

博士学位論文

不完全情報ゲームにおける
AI エージェントの戦略決定に関する研究

2020年3月

北海道科学大学大学院
工学研究科 工学専攻

高岡 勇樹

目次

目次	I
図目次	III
表目次	V
1 章 序論	1
1.1 ゲーム AI 概論	1
1.2 近年の動向と将来の展望	2
1.3 研究背景と新たなゲーム AI の提案	4
1.4 本論文の構成	7
参考文献	10
2 章 ゲームの分類と意思決定手法	11
2.1 ゲームの分類方法	11
2.2 完全情報ゲームと不完全情報ゲーム	12
2.3 ゲーム別の探索空間	13
2.4 ゲーム木とゲーム木探索	15
2.5 完全情報ゲームに用いられる手法	17
2.6 不完全情報ゲームに用いられる手法	23
2.7 機械学習	24
参考文献	25
3 章 ゲーム AI を強化する研究	26
3.1 完全情報ゲームの研究	26
3.2 不完全情報ゲームの研究	37
参考文献	40
4 章 本研究が対象とするゲーム	45
4.1 花札とは	45
4.2 花札の札種類	46
4.3 こいこいの遊戯方法	48
5 章 モンテカルロ木探索での強さの表現	53
5.1 モンテカルロ法の初期: 原始モンテカルロ法	53
5.2 原始モンテカルロ法の改良: モンテカルロ木探索	55
5.3 木探索と UCB の組み合わせ: UCT の誕生	56
5.4 UCB を利用した多腕バンディット問題の解法	58
5.5 UCT の適用	62
5.6 実験	66
5.7 まとめ	69
参考文献	71
6 章 ゲーム AI と面白さ	72
6.1 近年の動向	72
6.2 面白さを表現する研究の例	72
6.3 総括	74

参考文献.....	75
7 章 面白さを表現するためのアプローチ.....	76
7.1 研究背景.....	76
7.2 UCT プレイヤを改変した面白さの表現.....	77
7.3 花札のプレイヤへの面白さのアンケート.....	86
7.4 生体情報を用いた面白さの要素の絞り込み.....	96
7.5 まとめ.....	101
参考文献.....	103
8 章 結論.....	104
1 章 序論.....	104
2 章 ゲームの分類と意思決定手法.....	104
3 章 ゲーム AI を強化する研究.....	104
4 章 本研究が対象とするゲーム.....	105
5 章 モンテカルロ木探索での強さの表現.....	105
6 章 ゲーム AI と面白さ.....	105
7 章 面白さを表現するためのアプローチ.....	106
8 章 結論.....	106
今後の課題.....	107
業績一覧.....	109
投稿論文(査読あり).....	109
会議論文(査読あり).....	109
会議論文(査読なし).....	110

目次

図 2-1 三目並べの部分ゲーム木	15
図 2-2 ゲーム木の例	18
図 2-3 Min-Max 法: スコアを記録	18
図 2-4 Min-Max 法: 親ノードのスコアを計算	19
図 2-5 Min-Max 法: 探索終了時のスコア	19
図 2-6 α - β 法: β カット	21
図 2-7 α - β 法: α カット	21
図 3-1 チェッカー: 初期配置	27
図 3-2 チェッカー: 駒の移動	27
図 3-3 チェッカー: 駒の取得	27
図 3-4 チェスの駒	30
図 3-5 チェス: 初期配置	30
図 3-6 AlphaGo vs イ・セドル 第 4 局 黒 77 手目まで	35
図 3-7 AlphaGo vs イ・セドル 第 4 局 黒 101 手目まで	35
図 3-8 捨牌選択	37
図 3-9 副露選択	37
図 4-1 花札の札一覧	47
図 4-2 光札	47
図 4-3 タネ札	47
図 4-4 タン札	47
図 4-5 こいこい: 初期状態	51
図 4-6 こいこい: 札取り	51
図 4-7 こいこい: 取り札の整理	52
図 4-8 こいこい: 札が取れない場合	52
図 5-1 原始モンテカルロ法のモデル	54
図 5-2 モンテカルロ木探索: 有望な手に多くのプレイアウト	55
図 5-3 モンテカルロ木探索: 木の成長	55
図 5-4 多腕バンディット問題のイメージ	57
図 5-5 ある時点での UCB	59
図 5-6 UCB の更新	59
図 5-7 UCT 概要図	60
図 5-8 UCT: ルートを選択	61
図 5-9 UCT: 利得を加算	61
図 5-10 UCT: UCB を更新	61
図 5-11 探索する局面	64
図 5-12 相手の手札の推測	64
図 5-13 UCT 探索木	65
図 5-14 手札の決定	65
図 5-15 UCT プレイヤの獲得文数	66

図 5-16	ある局の場札と UCT プレイヤの手札	67
図 5-17	図 5-16 の局の開始時点	68
図 5-18	好ましくない札を選択してしまう局面	69
図 7-1	対局でやり取りされた得点	80
図 7-2	ある局の得点推移	81
図 7-3	ある局の UCB の変化	82
図 7-4	ある局の得点推移	84
図 7-5	プレイヤの年齢(N=61)	87
図 7-6	花札をやっている年数(N=61)	88
図 7-7	花札の面白いところ(N=61)	88
図 7-8	花札の面白くないこと(N=61)	89
図 7-9	花札の面白くないこと(N=61)	90
図 7-10	年齢層ごとの考える面白さ	91
図 7-11	年齢層ごとの考える面白くないこと	92
図 7-12	生体情報の例	96
図 7-13	実験用ゲーム画面	98
図 7-14	ゲーム中の 1 局面	99
図 7-15	計測された心拍数	99

表目次

表 2-1	ゲーム別の探索空間.....	13
表 4-1	花札の役一覧.....	50
表 5-1	図 5-16 の局面の UCB.....	67
表 5-2	図 5-17 の局面の UCB.....	68
表 7-1	対局者のコメント	85

1 章 序論

1.1 ゲーム AI 概論

ゲーム AI とは、本稿では「人間プレイヤーの相手をするコンピュータプレイヤーの意思決定に用いられる技術や、なんらかの意思決定手法を搭載したコンピュータプレイヤーそのもの」を指す。古くはチェスの AI であったり、パックマンの AI であったりと、特にゲームを一人でプレイする場合にはこれの良し悪しでプレイヤーからの評価が変動するものであり、ゲーム作成の中でも重要な部分である。従って現在では様々なゲームにゲーム AI が搭載されており、その種類は膨大なものになる。カードゲームの AI を 1 つ取り上げても、「ひたすら攻めることを考えるもの」・「相手の攻撃をとにかく耐えることを目標とするもの」・「特定のカードの組み合わせにより勝利を目指すもの」など、一つのゲームに複数の AI が搭載されているのも珍しくなくなっている。このことから、ゲームとゲーム AI はもはや切り離せないものとなっている。

1.2 近年の動向と将来の展望

近年はコンピュータの発展により、それまでゲーム AI の構築が困難とされてきたゲームの AI が構築できるようになった。それが囲碁である。囲碁は盤面が 19×19 と広く、それに伴って探索すべき空間が膨大に広い。よって、従来からある手法では対応できないことが多かった。これを解消したのが機械学習である。機械学習とは、大まかに言えば「機械自らが学習をして、様々なものを判別する」ものである。最近は多様な機械に機械学習が搭載され、一般の認知度が向上しているように感じる。例えば、自動で走行する掃除ロボット[1]や、自動ピックアップロボット[2]など幅広い場所で使われている。掃除ロボットでは、家具の配置・部屋の構造といったものを学習し、効率の良い掃除を実装している。またピックアップロボットでは、音声認識と組み合わせて好みのキャンディを伝えるとそれにあつたものを提供してくれる機能がある。かように機械学習は様々な分野で利用され、機械の目覚ましい発展に寄与している。この機械学習をゲーム AI に導入することにより、局面ごとのアルゴリズム作成が不要という利点が生まれ、膨大な過去のデータからゲームを学習させることによって局面に応じた適切な手を打つことのできるゲーム AI が作成できた。その結果、アマ四段程度だった棋力がもはやプロ棋士でも太刀打ちできないような力を持ったのである。

また、人間が直接コントローラを操作するゲームではないゲームの AI も研究されつつある。その一例が「汝は人狼なりや?」というゲームである。このゲームの進行はプレイヤー同士の会話によって行われ、コンピュータプレイヤーにはプレイヤーの会話を解釈することが求められる。そのためゲーム AI を作成するのが難しい部類であったが、自然言語処理の手法を取り入れることによりこれを実現してい

る。他にも定型コマンドを打ち込むことによりコミュニケーションを取る方式として、それを学習することでゲーム AI を強化することも行われている。いずれも大きな処理能力を必要とし、コンピュータが高性能になったことにより実現したものである。

1.3 研究背景と新たなゲーム AI の提案

本研究は特定のゲームの AI を「強化する」ことを目的としたものと、「面白くする」ことを目的としたもの、2つの観点から研究を行った。第一の「強くする」目的は、ゲーム AI の研究の基本は強くすることを目的としているからである。古くから現在までの研究を振り返ると、ゲーム AI の研究は強化することを主としている。であるから、本研究でも新たなゲームについてのゲーム AI を提案するべく、これまでに研究があまりなされてこなかったゲームである「花札」についてのゲーム AI を強くすることを目的とした研究を行った。花札とは日本に古くから存在するカードゲームの一種で、二人以上で行う遊戯である。花札は日本が主流のため、現在までに研究対象とされてこなかった。よって花札に従来のゲーム AI で用いられている手法を適用することにより花札の AI を強化することができるのではないかと考え、研究の第一段階として花札の AI を強くするという研究に着手することとした。

ゲーム AI を強化する研究が進行するにつれて、新たな問題点が浮上した。それは、「強すぎるゲーム AI は面白くないのではないか」という問題である。人間プレイヤーがゲームをプレイする際に重要なのは「自分が楽しいか」ということだと考えられるが、はたして自分が勝てないようなゲームをして楽しいと感じられるだろうか。普通ならば、あまりにも難しかったり、敵が強すぎたりする場合にはつまらないと感じるものではないかと考える。そこで、新たなゲーム AI の方向性として、「対戦をしていて面白いコンピュータプレイヤー」を構築できないかを思案した。この目的が達成された場合、よりゲーム AI の研究の幅が広がることが期待される。「面白いコンピュータプレイヤー」を作成するためには、「何が面白

いか」を特定する必要がある。言い換えれば、「花札の面白い要素はなにか」を求めることに等しい。仮に面白い要素が全て明らかになったならば、場面に合わせた面白さを表現することが可能であり、人間プレイヤーに楽しさを与えられるコンピュータプレイヤーが作成できるのではないだろうか。

面白さを与えるコンピュータプレイヤーの作成が可能であるかの検証として、実験の一つを実施した。その内容は「独自に定めた面白さに従って対局を行うコンピュータプレイヤーを作成し、人間プレイヤーとの対局を行って評価をする」ものであった。実験者に花札のプレイヤーを選び、対局をしてもらって評価点や問題点を聞き取ることを行った。その結果、初歩的ながらも面白さを与えることに成功したが、人間プレイヤーからの評価はあまり良くないものであった。

面白さを与えるコンピュータプレイヤーの作成は可能であると判明したものの、独自に定めた面白さでは不十分であり、面白さの要素を特定する必要に迫られた。

面白さの要素を特定するため、角度の異なるいくつかの実験を行った。一つは、花札のプレイヤーにアンケートをとり、プレイヤーが花札へどのような考えを持っているかを尋ねたものである。これにより得た結果を基に、面白さの分析を試みようとした。もう一つは、人間の持つ生体情報を利用した面白さの要素の特定についてである。生体情報には様々な種類があるが、ここでは心拍数を用いている。実験としては、ゲームをプレイ中の被験者へスマートウォッチを装着させ、心拍数の変化を観測した局面を分析することにより絞り込みを行えないか試みた。

これら2つの実験の結果より、新たなゲームAIの提案を行う。一つは「強い花札のAI」である。従来のゲームAIに適用されてきた手法を花札に応用したところ、人間プレイヤーを負かす程度の強さを表現することができた。よって構築した

ゲーム AI は十分な強さを備えており，新たなゲーム AI として採用できるものであると言える．もう一つは「プレイヤーに面白さを与える AI」である．こちらは現状では不完全であり，数々の問題を内包しているが，人間プレイヤーに面白さを与えるために必要な事項や，研究の発展のための足がかりを得ることができた．そこで，将来的な完成に向けてのアプローチを提案するものである．

1.4 本論文の構成

本論文の構成は以下の通りである。

1 章 序論

1 章ではゲーム AI に関する近年の動向を述べ、ゲーム AI の研究の現状を追う。その後行った研究の背景を記し、実験で得られた結果をもとに新たなゲーム AI の提案を行う。

2 章 ゲームの分類と意思決定手法

ゲームには数多くの種類があり、研究されてきたものは膨大な量になる。本章では第一にゲームを 2 つに分類し、それらの特徴を記述する。その後、ゲーム AI の研究で用いられる手法についてまとめる。

3 章 ゲーム AI を強化する研究

2 章で述べた分類に従い、それぞれの種類の先行研究を掲載する。ゲーム AI の研究は盛んに行われており、その種類はかなりのものである。本章ではそれらの内特に重要であると思われるものを取り上げることとする。

4 章 本研究が対象とするゲーム

本研究では、これまで研究されてこなかったゲームである「花札」を対象とした。本章では花札についての遊戯方法や、得点の計算方法などを述べる。

5 章 モンテカルロ木探索での強さの表現

ゲーム AI を構築する際にしばしば用いられるのが「モンテカルロ木探索」という手法である。モンテカルロ木探索はモンテカルロ法と木探索を混合させたもので、効率の良い木探索を実現している。本章ではモンテカルロ木探索について述べ、この手法が優れている点や、改良された木探索について示す。その後、花札に適用する場合の実装方法を記し、行った実験についてまとめる。

6 章 ゲーム AI と面白さ

近年、ゲームに関する研究は別の広がりを見せている。それは「人間プレイヤーに面白さを与えるゲーム AI」の構築に関する研究である。単純に強さを追い求めるのではなく、人間プレイヤーに楽しんでもらえるような AI を開発するようなものとなる。本章はそれらの研究について調査した結果を記載し、面白さを与える手法や実験の結果・人間プレイヤーの考えなどについて追う。

7 章 面白さを表現するためのアプローチ

本研究のもう一つの目標である、「プレイヤーに面白さを与える AI」を構築するために行った調査や実験結果について記述する。面白さを与えるためには、AI が「何が面白いのか」を知っている必要がある。そこで、研究の第一段階として花札の面白い要素を解明することとした。そのために花札のプレイヤーにアンケートを実施して面白さの要素を絞り込めないかを考えた。その結果、プレイヤーが花札に思うことなどが判明した。しかし、この手法はプレイヤーの主観によるところ

が大きく、それが正しいかどうかの議論がしにくい面があり、より定量的に面白さの評価ができる指標が必要であると考えた。

この面を解消させるべく、人間の生体情報を利用した特定の手法を考案した。生体情報とは、人間から発せられる種々の情報であり、代表的なものに心拍数や脳波、血圧がある。生体情報というものは得てして自らで制御しにくく、人間の心理状態がつぶさに捉えられるものだと考える。この生体情報を測定し分析することにより、対局をしているプレイヤーがどういう心理状態であるか、なにがそうさせているのかを解析できるのではないだろうか。この考えに基づき、対局者の生体情報を測定する実験を実施した。測定する情報は心拍数で、心拍数に反応がある局面の要素を解析することにより、花札の何が面白いのかを探る。それと同時に、今後同様の実験をする際に向けて、仮設が正しいのかを検証する作業も行う。

8 章 結論

この章では行った実験の結果をまとめ、結論として各章の総括を行う。合わせて今後行わなければならない課題を示す。

参考文献

[1] iRobot “i7 シリーズ | ロボット掃除機 ルンバ | アイロボット公式サイト”

<https://www.irobot-jp.com/product/i7/index.html>

[2] Google “Google I/O で注目の TensorFlow ロボット、その「賢さ」を支える
機械学習 | Google Cloud Blog”

[https://cloud.google.com/blog/ja/products/gcp/google-io-tensorflow-findyour
canfi](https://cloud.google.com/blog/ja/products/gcp/google-io-tensorflow-findyourcanfi)

2 章 ゲームの分類と意思決定手法

ゲームには複数の分類方法があるが、ここではゲームを「完全情報ゲーム」と「不完全情報ゲーム」に分類し、これらの特徴を述べる。その後、両者のゲーム AI がどのように発展してきたのかを示す。

2.1 ゲームの分類方法

本稿では様々な分類方法の内、プレイヤーが視認できる情報の有る無しにより分類する手法を取り上げる。ゲームをプレイする際には、プレイヤーは種々の情報を参照して進行させることがもっぱらである。例えば、将棋で言えば盤上の駒の配置や取得された駒の種類・数、駒を動かした後の状況がどう変化するか、などであり、麻雀で言えば自分の手牌 14 枚の構成や得点の期待値、各家の持ち点などである。これらの情報を逐次参照しながらプレイヤーはゲームを進行させ、いかにして勝つかを考えるわけである。この時、ゲームにより「全ての情報が参照できる」ものと、「プレイヤーからは見えない情報が存在する」ものの二つに分類ができる。先述の将棋を取り上げると、盤上の状況や取った駒というのは全てのプレイヤーが参照可能であり、隠されている情報は何も存在しない。一方、麻雀においては、壁牌は自分が自摸ったものしか分からず、同時に他家の手牌構成や裏ドラといったものは誰も知る術がないのである。このような「参照できない情報の有る無し」に着目し、その情報が無いものを「完全情報ゲーム」、有るものを「不完全情報ゲーム」と分類し、両者の特徴を記述することとする。

2.2 完全情報ゲームと不完全情報ゲーム

完全情報ゲームには、囲碁や将棋，チェス，チェッカーなどが該当する．いずれのゲームにおいても，全ての情報が参照可能であり，プレイヤーは何十手もの先読みにより勝利を目指す．これらのゲームについては，古くから研究されてきた歴史がある．また共通することがらとして，コンピュータの発展とともに AI の実力が向上している点がある．

不完全情報ゲームには，麻雀やトレーディングカードゲームの多く，ポーカーやブラックジャックといった，比較的賭け事に使われやすいゲームが該当する．これらのゲームは相手の手札をそれまでの行動から予測し，自分が勝てるようにプレイをするタイプのゲームである．このタイプのゲームは近年研究が活発であり，その実力は千差万別になる．採用する手法も様々であり，今後新たな手法の開発とともに実力の向上が期待される分野である．

2.3 ゲーム別の探索空間

ゲームにおける探索空間の数は「そのゲームで出現する全ての局面の数」である。この数が大きくなるに従って探索が難しくなるものである。これをまとめたものが表 2-1 である[1]。

表 2-1 ゲーム別の探索空間

ゲーム	探索空間
チェッカー	10^{20}
オセロ	10^{28}
チェス	10^{50}
将棋	10^{71}
囲碁(9 路盤)	10^{38}
囲碁(19 路盤)	10^{171}

表 2-1 は完全情報ゲームでの探索空間であり、不完全情報ゲームにおいてはより多くの探索空間がある。両者に言えることは、広大な空間を全て探索するのはほぼ不可能であり、通常の探索では深さを決めて探索を行うのが一般的である。ここで、勝負の結果が出ていなければどちらが有利なのかを判定する指標が必要となる。この指標を評価関数と称し、評価関数の良し悪しも探索を行う上で重要なものである。なぜなら、探索で得られた優劣を誤認すると間違った手を選択する可能性があるためである。仮に探索結果が有利な局面を得たとしても、評価の段階で不利と判断すればその手を打たなくなってしまうかもしれない。そして逆

に不利な局面につながる手を有利と見て選択するかもしれない。その点において、評価関数の設計は重要であり、難しいものである。

2.4 ゲーム木とゲーム木探索

探索空間が狭いゲームにおいては、ゲーム木を作成して探索を行うと強い AI が得られることが多い。ゲーム木とは「プレイヤーが取る行動を木構造で表したグラフ」である。これを辿って適切な手を選択するアルゴリズムが木探索である。また、開局から終局までの全ての手を正確に作成すると完全ゲーム木、ある地点から途中までを作成すると部分ゲーム木と呼ぶ場合もある。ゲーム木では、手を打った局面が各ノードになり、枝がそのノードにおいて打つことのできる手である。子ノードを持たないノードはゲームが終局した局面であることを示す。本稿では、親ノードを持たないノードをルートノード、子ノードを持たないノードを葉ノードと呼称する。

例として三目並べの二手目までの部分ゲーム木を図 1 に示す。ここでは、回転させて同じ形状になる手は省略している。図 2-1 のように手を展開させ、行動に合わせて木を辿ると最適な手が見つかる。葉ノードに到達したならば、勝敗を判断して再び探索を行う。最終的にどの手を選択すべきかを探索結果より求め、ゲームを進行させる。この「手を辿って行動を評価する」アルゴリズムがゲーム木探索である。

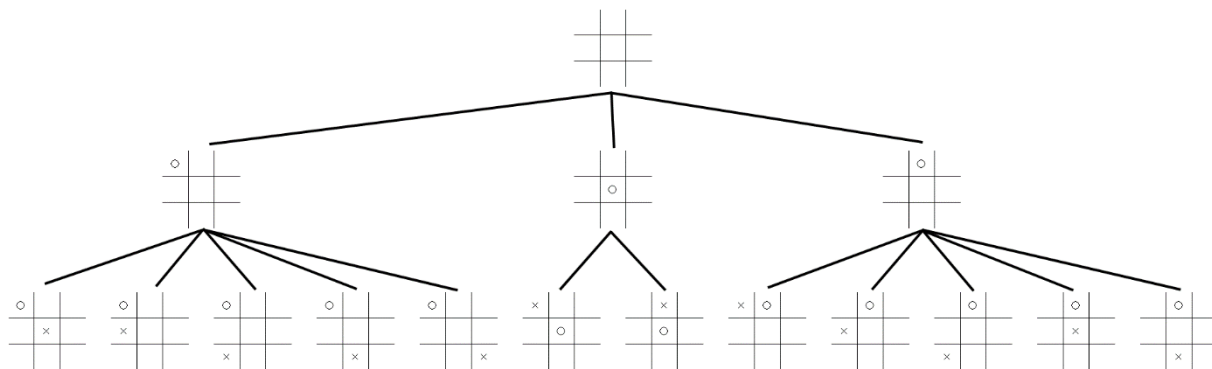


図 2-1 三目並べの部分ゲーム木

完全ゲーム木が作成できるならば、そのゲームの結果がどのようなになるかを解析することが可能となる。例えば、三目並べにおいては両者が最善手を選択すると必ず引き分けとなることが知られているし、これを拡張した五目並べでは先手が勝つと判明している。よってゲーム AI の実力もかなりのものとなり、強い AI が数多く作成されている。

2.5 完全情報ゲームに用いられる手法

完全情報ゲームでは 2.4 のゲーム木探索を用いることがしばしばある。この時に利用されるのが「Min-Max 法」や「 α - β 法」であり、これにより効率的な探索を行っている。

2.5.1 Min-Max 法

「Min-Max 法」とは、「互いのプレイヤーが最善を尽くすと仮定した上で、自分にとって最も有利になる手を選ぶ」手法のことである。この手法は 1928 年に数学者ジョン・フォン・ノイマンにより提唱されたのが最初であるとされる [2]。

例示のため二人でプレイするゲームを想定する。このゲームは対局のスコアが数値として現れるゲームで、先手のプレイヤーが有利な場合はスコアが大きく、後手が有利な場合はスコアが小さく現れると仮定する。これより、先手はスコアを大きくすべくプレイし、後手は小さくすべくプレイすることを目指すだろう。これを探索するときのゲーム木の例が図 2-2 である。図 2-2 のゲーム木は深さが 4 で、赤のノードが先手の手番である局面、青のノードが後手の手番である局面、黒のノードは終局、丸囲み数字は対局結果の数値である。探索は全てのノードをたどることから始まる。ルートノードより木をたどり、葉ノードまで到達したらスコアを記録し、木に記す(図 2-3)。親ノードのスコアはこれを用いて付けられる。深さ 3 のノードは子ノードが 1 つだけであるので、子ノードのスコアがそのまま適用される。子ノードが複数ある場合は、先手ならば子ノードのうち最大のスコアが自身のスコア、後手ならば子ノードのうち最小のスコアが自身のスコアである(図 2-4)。これは、先手はスコアを大きくすべくプレイし、後手は小さく

すべくプレイするところより来ている。このスコア付けをルートノードに達するまで繰り返す。全てのノードのスコア付けが終了したならばゲームの特性が見えてくる。ここで挙げたゲームはスコアが 50 という結果になる(図 2-5)。よってここで考えたゲームは先手・後手ともにスコア 50 になるような手を打ち、ゲームが終局するだろう。以上の手順で打つ手を決定するのが **Min-Max** 法である。

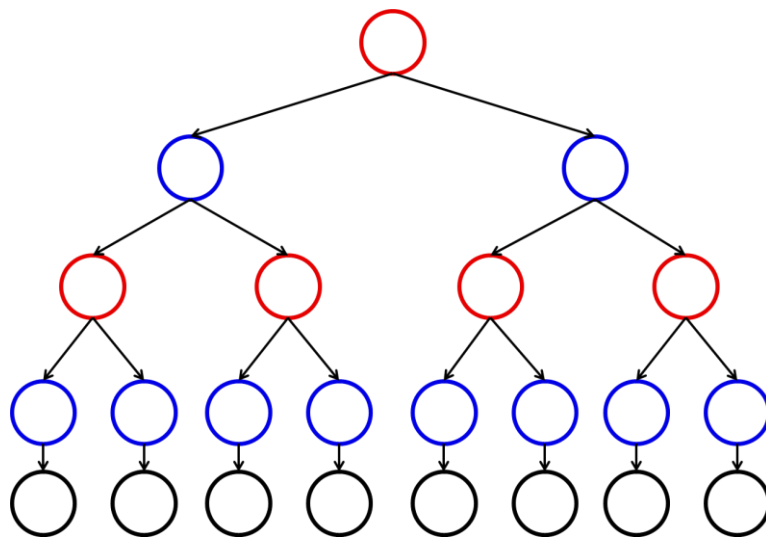


図 2-2 ゲーム木の例

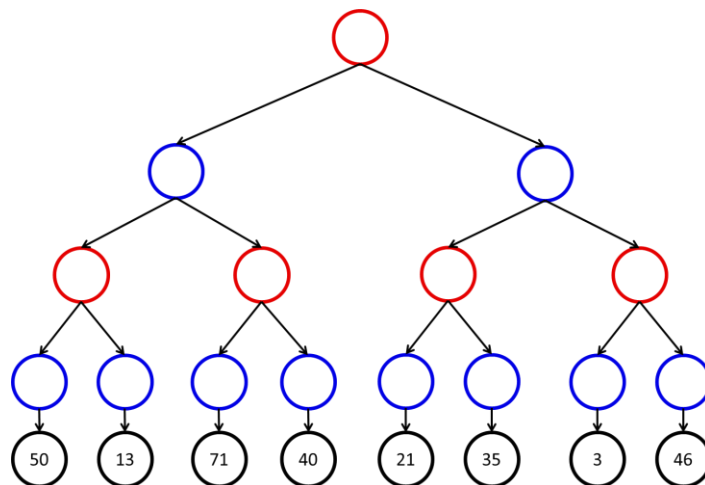


図 2-3 Min-Max 法: スコアを記録

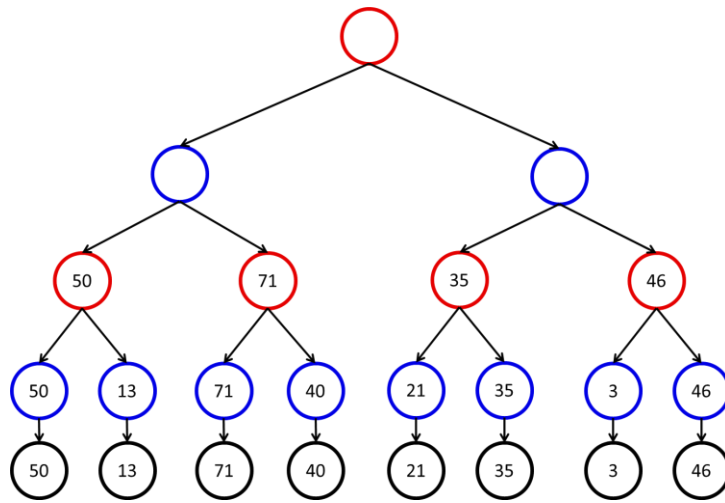


図 2-4 Min-Max 法: 親ノードのスコアを計算

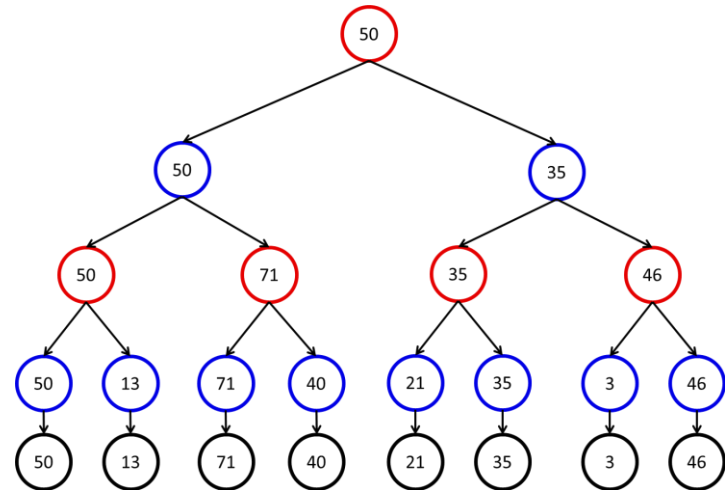


図 2-5 Min-Max 法: 探索終了時のスコア

2.5.2 α - β 法

α - β 法は Min-Max 法に枝刈りの概念を付与した木探索アルゴリズムである。これは 1975 年、ドナルド・クヌースがチェスで経験的に使われている手法をまとめ、Min-Max 法の改良として開発したのが始まりである。この中でクヌースは、葉ノードが n 個ある時、 α - β 法が最も効果がある場合で \sqrt{n} 個のノードを探索すれば良く、木探索の効率化が図れることを明らかにした[3]。

ゲーム木が図 2-3 の状態であるとし、左の子ノードから探索するとルールを定める。図 2-6 の A のノードに着目すると、スコア 71 の子ノードを読んだ段階で自身のスコアは 71 以上が決まる。なぜなら A のノードは先手の局面で、そのスコアは子ノードの最大値になるからである。A の親ノードは最小値がスコアとなるため、A のノードからは選択されず以降の探索を行ってもスコア付けに影響を及ぼさず処理時間のみが消費される。よって A の右の子ノードは枝刈りされ効率化を行う。これを β カットと呼ぶ。 β カットはスコアがある値以上となる場合に発生する。次に図 2-7 の B のノードに着目する。B の親ノード C は後手の局面で、スコアは子ノードの最小値を取る。これより C のスコアは B を読んだ時点で 35 以下が確定する。ルートノードは子ノードの最大値を取るので C 以下からは選択されず、同様に以降の探索は影響がなく枝刈りがされる。これを α カットと呼ぶ。 α カットはスコアがある値以下となる場合に発生する。このような α - β 法により木探索が効率化し、探索時間を大幅に減少させることができる。これ以降のゲーム木探索アルゴリズムは α - β 法を採用しているものが多い。

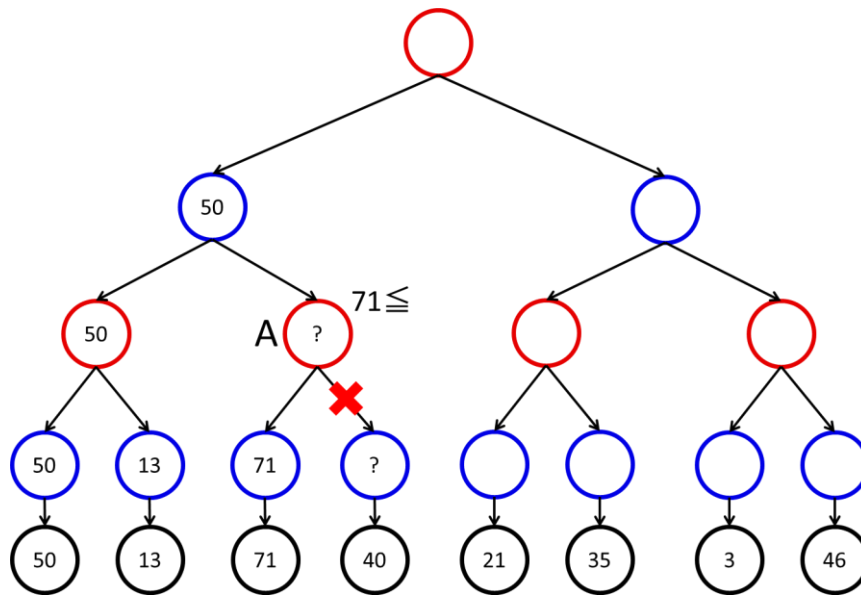


図 2-6 α - β 法: β カット

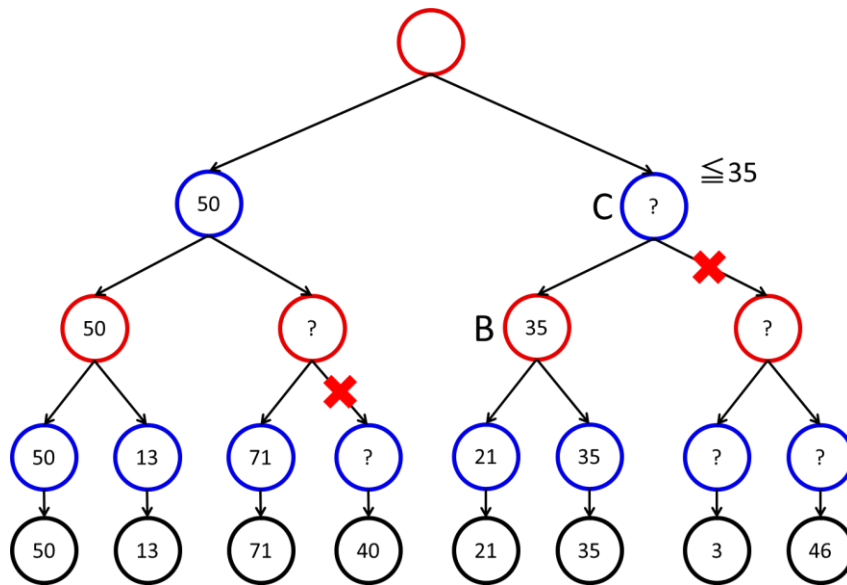


図 2-7 α - β 法: α カット

完全ゲーム木の作成できない広大な探索空間を持つゲームの場合では、モンテカルロ木探索が利用されることがしばしばある。これは木探索を途中で打ち切った場合でもある程度正確な評価値を得ることのできる手法であり、評価関数の設計が不要という利点がある。この手法が初期の囲碁・将棋 AI に採用されて実力が向上した事例から広く利用されるようになったもので、複雑なゲームではモンテカルロ木探索やその改良である UCT(後述)が現在でも採用されることのある手法である。

2.6 不完全情報ゲームに用いられる手法

不完全情報ゲームは、各々の特性により取るべき戦略や攻略法が変わってくるものであり、ある手法を別のゲームに搭載してもうまくいかないことが多い。そのためゲーム毎に違う技術が必要で、多数のゲームに採用されているような代表的なものを述べるのは困難である。不完全情報ゲームの一種である「汝は人狼なりや?」を例にとると、このゲームは会話をしながら進行させるため、人間が発言した内容の解析に自然言語処理の技術が要求される。次にそこから得た情報を基に自らのアクションを決定するが、これには他のプレイヤーの発言の真偽を見抜く技術・信頼を得てプレイヤーを説得するための行動選択など、従来にはなかった要素を勘案しなければならない[4]。そして、自然言語処理は発言を行わないゲーム、例えばセブンブリッジには適用する必要がなく、他のプレイヤーとの協調もない。このように、あるゲームに特化した技術を他分野へ応用することは困難であり、そのため統一された手法を例示するのは難しい。これまでの研究では、完全情報ゲームと同様にモンテカルロ木探索を採用したものが多くあり、ある程度の結果を残している。しかし近年になって、完全情報ゲーム・不完全情報ゲーム共に機械学習を導入するケースが増加している。その点で機械学習が主流になりつつあると言えるかも知れない。

2.7 機械学習

ゲームにおける機械学習では、人工ニューラルネットワーク(ANN と呼ばれる)を取り入れたものが多くある。とりわけ、ANN を多層にしたディープニューラルネットワークによるディープラーニングによるゲーム AI が顕著な実力を見せている。将棋の AI である **Ponanza Chainer** はディープラーニングのライブラリである **Chainer** を学習に採用し[5]、2017 年の第 27 回世界コンピュータ将棋選手権で準優勝を達成している[6]。また対戦型ゲームに留まらず、一人用ゲームのプレイにも機械学習を利用する例が存在する。**Hado** らは様々な一人用ゲームにディープラーニングと Q 学習を組み合わせた **DQN(deep Q-network)**を適用して、人間プレイヤーのスコアをどの程度上回れるかを調査している[7]。その結果、ビデオピンボールでは人間プレイヤーとのスコア比で最大 7220%向上、ブロック崩しでは同 1397%向上というものが得られたとしている。よって、機械学習は一人用ゲームでも有用であり、ゲームの研究、もっと言えば完全情報ゲーム、不完全情報ゲーム、一人用ゲームといった分野に関係なく機械学習を用いた研究がもっと盛んになされるのではないかと考える。

参考文献

- [1] 副島佑介: コンピュータ囲碁における Root 並列化について, 東京工業大学
修士論文(2010)
- [2] Tom Siegfried: A beautiful math: John Nash, game theory, and the
modern quest for a code of nature, Natl Academy Pr(2006)
- [3] Donald E. Knuth, Ronald W. Moore: An analysis of alpha–beta
pruning., Artificial Intelligence., Vol.6, Issue 4, pp.293–326(1975)
- [4] 鳥海不二夫 “第 1 回 「人狼知能」 って何?”
<https://book.mynavi.jp/manatee/detail/id=76458>
- [5] 山本一成, 下山晃ら: 第 27 回世界コンピュータ将棋選手権 Ponanza Chai
ner アピール文章
- [6] 世界コンピュータ将棋選手権 “第 27 回世界コンピュータ将棋選手権”
<http://www2.computer-shogi.org/wcsc27/>
- [7] Hado van Hasselt, Arthur Guez and David Silver: Deep Reinforceme
nt Learning with Double Q-learning, ArXiv e-prints, arXiv:1509.0646
1v3(2015)

3 章 ゲーム AI を強化する研究

前章で述べた手法や新たなアルゴリズムにより構築されたゲーム AI を本章で示す。いずれのゲームも探索空間が広大でゲーム AI の構築が困難であるとされたものである。

3.1 完全情報ゲームの研究

3.1.1 チェッカー

チェッカーは「敵駒を全滅させると勝利」するゲームである。日本においてはあまり知られていないが、世界的にはポピュラーなゲームだとされる。通常のチェッカーより広い「チェッカー・ドラフト」は世界 4,000 万のプレイヤーがいるとされている[1]。

チェッカーの初期盤面を図 3-1 に示す。黒を先手とし、各プレイヤーは手番で駒を 1 つ動かす。その際、以下のルールに従う。

- 通常の駒は斜め前にのみ動ける(図 3-2 左)。
- 敵陣一番奥まで到達すると駒が昇格する。この時、その駒をキングと呼称する。キングは斜め後ろにも動ける(図 3-2 右)。
- 敵駒が進行方向にあり、その次のマスが開いている場合は敵駒を取らなくてはならない(図 3-3)。動いた先でも取れる場合は連続して取る。取れるのに他の駒を動かすのはルール上できない。
- 移動を繰り返して敵駒を全滅させると勝利、動かせる駒がなくなると敗北である。

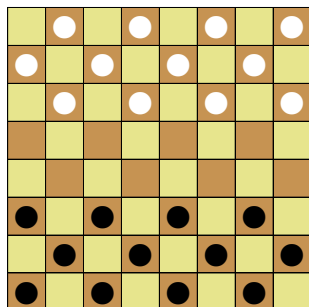


図 3-1 チェッカー：初期配置

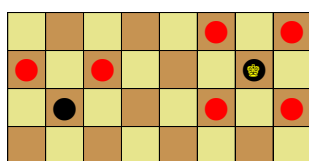


図 3-2 チェッカー：駒の移動

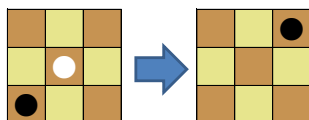


図 3-3 チェッカー：駒の取得

チェッカーの局面を評価する要素としては「駒の数、キングの数、キングになれる可能性のある駒の数」などが挙げられる。まず、全滅すると敗北であるから、自軍の駒がいくつ残っているか、相手との駒差はどのくらいかという要素は重要である。次にキングの枚数がどの程度多くあるかを考慮する。キングに昇格すると駒の自由度が向上する。キングは攻めにも守りにも有用で、大切な駒である。その他、進んだ先の陣形がどのように変化するかを先読みし、自軍の駒を塞がないようにしなければならない。チェッカーは駒を一つ動かすのが行動の全てである。これより AI はどの駒を動かすと勝利に近づくのかを考えるものである。

初期の AI はアーサー・サミュエル(Arthur Lee Samuel)が作成した「Samuel Checkers-playing Program」が世界初のチェッカープログラムとして知られてい

る [2]. サミュエルは自身に勝てるチェッカーのプログラムを作りたいと考え、どうすれば強くなるのかを追求した。

サミュエルの開発したプログラムは、探索木を用いて局面を探索し探索結果から差し手を評価してゲームの進行を行うものである。木探索中に選択する次の一手は Min-Max 法により決定する [3]. その際、 α - β 法を適用して探索木の枝刈りを行い、探索の効率化を図っている [4].

またサミュエルは機械学習の手法も取り入れた。プログラムは自分自身と数千回も対局し、良い局面・悪い局面がどんなものか自ら学習した。結果この手法は上手く行き、1962年にコネチカット州チャンピオンを倒すことに成功した [5].

その後の 1989 年には、アルバータ大学のジョナサン・シェーファー (Jonathan Schaeffer) らによって「Chinook」のプロジェクトが開始された [6]. Chinook プロジェクトの目標は当時最も強いチェッカープレイヤーであるマリオン・ティンズリー (Marion Tinsley) を打ち負かすことであった。ティンズリーは Chinook と対局するまでの 40 年のプロ期間中、たったの 3 敗しかないようなプレイヤーである。

Chinook は 1990 年、ティンズリーに次ぐ成績を収めティンズリーへのタイトル挑戦権を得ることに成功した。1992 年に対局が実現したが、この時は 2-4-33 で Chinook が敗れている [7]. それでも Chinook の 2 勝は、ティンズリーにつけた 4 つ目および 5 つ目の黒星ということで大変に意味のある対局だった。後の 1994 年に再び対局したが、6 戦全て引き分けに終わった。ティンズリーはこの後、健康上の理由により棄権し [7], Chinook が世界チャンピオンとなった。

シェーファーはさらにチェッカーの解析を続け、2007 年に「初期局面から両プレイヤーが最善を尽くした場合、いかなる場合でも引き分けに至る」と結論付けた

[8]. これにより、Chinook は引き分けることはあっても負けることが無くなり、人間が勝利することは不可能である。

最終的に引き分けを証明したプログラムは木探索アルゴリズムと終盤データベースよりなる[6]. 終盤データベースは、事前に計算したチェッカーの終盤局面の評価値をデータベース化したものである。このデータベースの構築で木探索アルゴリズムの大幅な効率化が可能となり、木探索アルゴリズムの組み合わせによりチェッカーが完全に解析された。

3.1.2 チェス

チェスの研究は長い歴史がある。ここでは、ある程度重要であると思われるものを記述する。

チェスとは「敵駒を取りながら敵のキングを追い詰めると勝利」するゲームである。日本の将棋に似るが、取った駒を使えない、成駒の概念がほぼない、といった違いがある。

チェスの駒には6種類あり、それぞれ「キング・クイーン・ルーク・ビショップ・ナイト・ポーン」である(図 3-4 左より)。駒の動かし方も将棋と似通っており、大まかに言えば

- キング: 王将
- クイーン: 飛車と角行を合わせたもの
- ルーク: 飛車
- ビショップ: 角行
- ナイト: 桂馬を上下左右の方向

- ポーン: 歩兵

である。駒の取得もほぼ同様に、移動先にある敵駒を取ることができる。他、より複雑なルールがあるがここでは割愛する。

チェスの初期配置を図 3-5 に示す。ここから白を先手として始め、相手のキングを詰ませることを目的とする。ここで、強い AI とはキングを効率よく詰ませるものであり、そのためにどの駒を移動させるか思考するものである。



図 3-4 チェスの駒

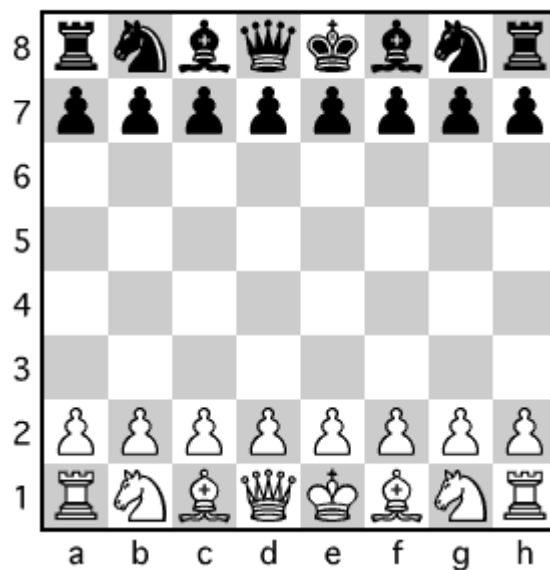


図 3-5 チェス: 初期配置

初期の AI にディープ・ブルーが挙げられる。ディープ・ブルーはチェス専用のスーパーコンピューターであり、前身のディープ・ソートを改良したものである。これは 1990 年代に IBM で開発された。ディープ・ブルーは計算速度に優れている特徴がある。これは、30 プロセッサの IBM PC と 480 個のチェス局面探索エ

ンジンからなり、各プロセッサにも 8 個の探索チップを搭載している。このスペックにより、1 チップだけで秒間 200 万以上の局面の探索を可能とした[9]。

ディープ・ブルーもまた、木探索により局面を探索する。その際、全探索をするのではなく、重要な局面を深く読む選択的探索を行っている。探索中に出現した重要な局面にリソースを割り当て、それを重点的に、さらに深く読む。選択的探索と他に実装されている手法によって、ディープ・ブルーは世界クラスレベルのチェスを指すことができるようになった。

時代が下るとコンピュータが安価になり、また性能も向上した。それによってチェス専用のスーパーコンピューターでなくても、一般のコンピュータに AI を搭載できるほどにまでなった。顕著な例がスマートフォンにも搭載できる Pocket Fritz 4 である。これはスマートフォン上の AI でありながら 2009 年のチェス上位者が集まる大会で優勝し、レーティング 2898 を達成した[10]。

さらに時代が下り、AI の棋力向上によって人間プレイヤーとの対局が実施されるようになった。古くはディープ・ソートが 1989 年に当時の世界チャンピオンであるガルリ・カスパロフ(Garry Kimovich Kasparov)と対戦をしているが、このときは 2 回戦を 0-2 で敗れている[11]。その後の 1996 年にディープ・ブルーとの対戦が行われ、ディープ・ブルー側が 1 勝を挙げたことで話題となったが、結果的には 6 番勝負を 1-3-2 で敗れている[12][13]。しかし、翌年の 6 番勝負は 2-1-3 でディープ・ブルーの勝ちであった。対局後に「コンピュータが世界チャンピオンに勝利した」と世界中へ大変なニュースが流れ[14]、ディープ・ブルーの強さを世界へ知らしめた。その後もコンピュータと人間の対局は続き、ほぼコンピュータの勝ちで占められている。2016 年にはグランドマスターのセルゲイ・エレン

ブルグ(Sergey Erenburg)と AI の Komodo との対局が行われ、Komodo 側のポーン 2 駒落ちでの対局だったにも関わらず人間の 2 敗 2 分と、ハンデがあっても勝てない状況になっている [15].

3.1.3 囲碁

囲碁は日本では馴染み深いゲームであり、AI の研究が盛んに行われているゲームでもある。表 2-1 の通り探索空間が膨大で、さらに評価関数が作成しにくいという問題もあり AI を構築するのは不可能であると考えられてきた。評価関数の設計が困難であるのはいくつかの要因があるが、一つは将棋やチェス・チェッカーのように駒に優劣がなく、駒得のようなわかりやすい指標がないこと、もう一つは局所的な得が最終的な得とならないことが多々あることであると言われている [16]. よって囲碁には評価関数を用いる方法以外のアルゴリズムが求められる。そこで注目されているのがモンテカルロ木探索である。モンテカルロ木探索を採用した囲碁 AI として Crazy Stone [17] を挙げる。Crazy Stone はモンテカルロ木探索によって打ち手を決定するが、プレイアウトの際に完全ランダムで行うのではなく、囲碁特有のパターンを用いて精度の向上を図っている。このパターンによって最善手が発見されたならば、通常の UCT よりも優遇して選択するようになっている。Crazy Stone は 2006 年のコンピュータ・オリンピックの囲碁 9 路盤で優勝し、強いプログラムであることを証明してみせた [18]. その後、他のプログラムがモンテカルロ木探索を取り入れるようになり、今では囲碁 AI の主流となっている。

そして現在急激に棋力を向上させているのが Google DeepMind の AlphaGo である。AlphaGo では探索にディープラーニングとモンテカルロ木探索を用いて効率よく探索を行っている [19].

AlphaGo はまず評価関数をディープラーニングによって学習する。評価関数は 2 つあり、それぞれポリシー関数とバリュー関数と呼ばれる。ポリシー関数は盤面のどこに着手すべきかの確率、バリュー関数は盤面の評価値を導出する。この 2 つをディープニューラルネットワークの教師あり学習で学習させる。教師データは、ポリシー関数は 3000 万局面以上の盤面データ、バリュー関数はポリシー関数で作った盤面とその盤面での勝率を用いている。ここで特徴的なのが、ポリシー関数は盤面を画像のように学習させることである。もともとディープラーニングのモデルの一つである畳み込みニューラルネットワークは画像認識分野で大きな成功を収めていて、これをゲームに応用し、盤面の画像を教師データとして学習をさせるのである。こうして学習された評価関数とモンテカルロ木探索を組み合わせ、対局中に打つべき手を決定する。モンテカルロ木探索は、木を辿る部分にポリシー関数を使って「ありそうな手をより深く」読み、バリュー関数を使って「プレイアウト時の勝率を補完」することを行う。以上が AlphaGo の根幹である [20][21].

他のゲームに違わず、AlphaGo も人間プレイヤーとの対局を実施している。

AlphaGo が対局したもので特に有名なのが 2016 年に行われた韓国のプロ棋士であるイ・セドル九段とのマッチである。イ・セドルとは、21 年のプロ期間に世界王者を 18 回獲得しているような人物で、事前記者会見では自信を見せていた [22]. しかし、結果は 4-1 で AlphaGo の勝利だった。

コンピュータ対人間での印象的な対局として、この対局の第4局を挙げる。この対局はイ・セドルが勝利したが、コンピュータの弱点が見えた対局でもあった。以下の棋譜は[23]より引用した。図 23 は第4局 77 手目までを表示したものである。ここままで黒の AlphaGo が優勢で、白のイ・セドルは苦戦を強いられていた。しかし、イ・セドルの 78 手目、白 11 九が妙手となり、AlphaGo が一転不利になった[24]。AlphaGo はこの後悪手を連発し、黒 101 手目までそれが続いた(図 24)。結局 AlphaGo はこの差を埋められずに 180 手で投了した。この局はモンテカルロ木探索の弱点であると David Ormerod は評している。David によれば、白 78 のような局面における妙手が 1 つしかない場合、モンテカルロ木探索はそれを考慮せずに見落としやすくなると述べている。これは手筋を成立させる順序がただ 1 つであることがしばしばあるために見落とす可能性が高くなるとしている[25]。

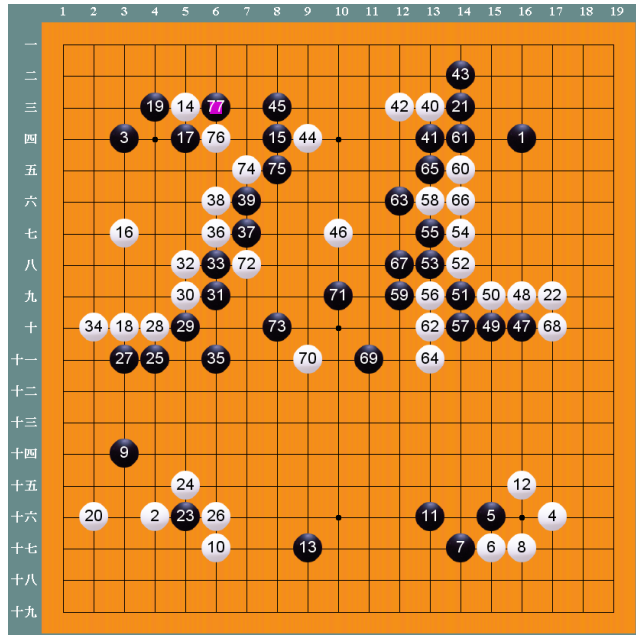


図 3-6 AlphaGo vs イ・セドル 第4局 黒77手目まで

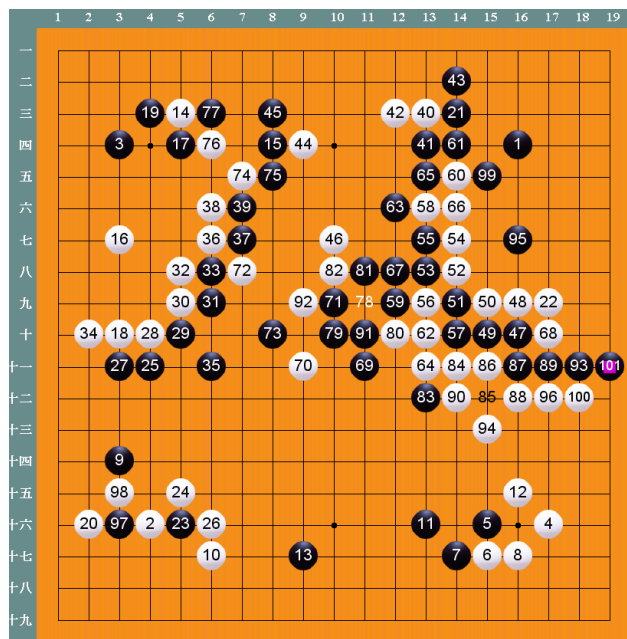


図 3-7 AlphaGo vs イ・セドル 第4局 黒101手目まで

現在 AlphaGo は AlphaGo Zero を経て、汎用 AI の AlphaZero となっている [26]. AlphaGo Zero は自己対局、ニューラルネットワークの学習と評価により成り立っており、AlphaGo との対局では AlphaGo に対してレーティングで 327 上回る性能を示したとしている。AlphaZero はこれを一般化したもので、AlphaZero を利用した開発したチェスと将棋 AI は世界チャンピオンであるプログラムを破った [27].

3.1.4 総括

これらの研究より、多くの完全情報ゲームはいずれも高い実力を備えており、人間プレイヤーが勝つことは難しくなっている。またこれ以外の将棋 [28] やリバーシ [29]、連珠 [30] においても顕著な成績を挙げ、すでに人間プレイヤーでは勝利できない、あるいはいずれそうなることが予測される。よって、現在では完全情報ゲームはおおよそどのゲームでも強い AI が構築できると言える。

3.2 不完全情報ゲームの研究

一方、不完全情報ゲームにおいては AI の実力は向上しつつある段階であり、研究の余地がある分野である。またそれまでゲーム AI の作成が困難であるとされてきたものも出現しており、研究が盛んになされている状況である。

3.2.1 麻雀

麻雀の AI は三木らが行った研究が挙げられる [31]。三木らは麻雀の捨牌選択(図 3-8)や副露選択(図 3-9)に UCT を採用し、上級者の牌譜一致率や対局の実力を調査したものである。

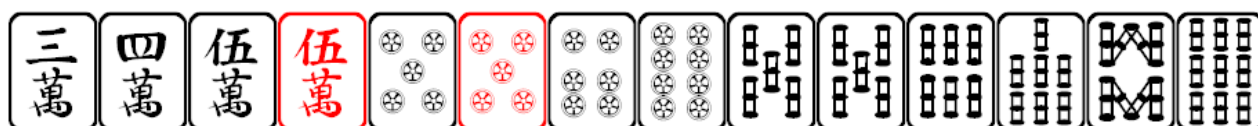


図 3-8 捨牌選択

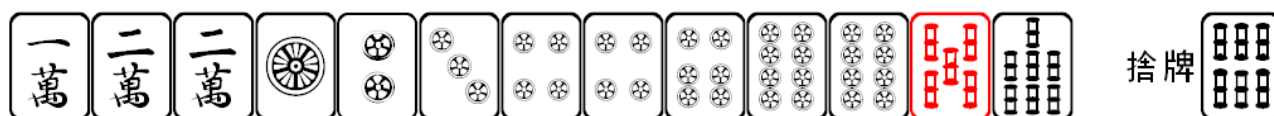


図 3-9 副露選択

まず牌譜一致率については、UCT のパラメータを変化させた場合に一致率がどう変わるかを調査している。変化させるパラメータは

- 報酬(\bar{X}_l)
- 式(5-1)中の c
- UCT ループ回数

である。報酬には、得点収支・順位点・得点収支の正規化・順位点の正規化などを適用する。

実験の結果として、報酬に関しては得点収支を報酬としたときに最も一致率が高いとしている。cは1000のときが最もよく、得点のスケールに合っているからであると分析している。ループ回数の変化は、一部例外はあれども多くなれば良いとしている。

対局実験では、グリーディプレイヤーとUCTプレイヤーの対局、SVMプレイヤーとUCTプレイヤーとの対局、この2種類を設定している。対グリーディプレイヤーの対局結果は、UCTプレイヤーが和了率と得点収支のいずれも上回り、対SVMプレイヤーでもほぼ互角の性能を発揮したとしている。そして結論として、UCTは麻雀のような複雑なゲームにも有効である可能性があるとして結論づけた。

また、AlphaGoの手法を用いたものに、Dwango Media Villageの開発した深層学習麻雀AIのNAGAがある[32]。NAGAは4つのCNNから構成されていて、自分から見える情報を入力として牌選択を行う。学習には、オンライン麻雀である天鳳の上級者の牌譜を用いている。NAGAは天鳳での対局を繰り返した結果、天鳳の段位で八段を記録したとしている。これより、CNNは完全情報ゲームだけではなく不完全情報ゲームにも利用できる可能性があるといえる。

3.2.2 汝は人狼なりや?

このゲームは他のゲームとは異なる要素により困難であるとされている。汝は人狼なりや?(以下「人狼」とする)は、プレイヤーとの対話によりプレイを進行させるゲームである。大まかに言えば、プレイヤーは村人陣営と人狼陣営に別れて村人が全滅する前に人狼を全滅させるのが目的となる。この際、プレイヤーの内誰をゲームから外すかを協議により決定するため、発言内容を分析したり、真偽を判定

したりしなければならない。よって、チャット形式で進む場合は自然言語処理の領域が関係してくる。また行動を定型化したプロトコル形式で進む場合でも、より複雑な学習が求められる。この人狼の AI を発展させることを目的とした「人狼知能プロジェクト」が発足している[33]。プロジェクト内では競技会やワークショップを通じて実力の向上を図ろうとしており、今後が期待される分野である。その他、SVM を用いた人狼の推定[34]が報告されている。

3.2.3 総括

不完全情報ゲームの研究は道半ばであり、研究の発展の余地が大いにある分野である。またゲームごとに要求される技術も多種多様であり、ゲームに合った手法を選択する必要がある。このような現状の中、完全情報ゲームに利用されているモンテカルロ木探索や改良の UCT を用いた AI、そして機械学習を用いた AI が出現してきており、研究が活発になされているのが現状である。今後はより高度な学習方法や新たな手法が出現し、より大きく研究が進展するものではないかと考える。

参考文献

- [1] 日本チェッカー・ドラフツ協会(JCDA):
<https://sites.google.com/site/japancheckersdraughts/>
- [2] Gio Wiederhold and John McCarthy and Ed Feigenbaum: Memorial Resolution: Arthur L. Samuel, Stanford University Historical Society.

(1990)
- [3] Arthur Samuel: Some Studies in Machine Learning Using the Game of Checkers, IBM Journal July 1959 (1959)
- [4] Richard S. Sutton and Andrew G. Barto: Reinforcement Learning: An Introduction, MIT Press, pp.267-269(1990)
- [5] Jeremy Howard: The wonderful and terrifying implications of computers that can learn, TED(2014)
- [6] 岸本章宏: チェッカー解明秘話, 情報処理学会誌, Vol.48, No.11, pp.1257-1263(2007)
- [7] Jonathan Schaeffer: Marion Tinsley: Human Perfection at Checkers? Department of Computing Science, University of Alberta(発行年不明)
- [8] Jonathan Schaeffer and Neil Burch and Yngvi Björnsson et al.: Checkers Is Solved, Science, Vol.317, pp.1518-1522(2007).
- [9] Murray Campbell, 丸山宏(翻訳): 人工知能におけるゲームの役割
情報処理学会誌, Vol.44, No.11, pp.1152-1156(2003)

- [10] The Week in Chess 771
<https://web.archive.org/web/20110930232108/http://www.chess.co.uk/twic/twic771.html#13>
- [11] David N. L. Levy and M. Newborn: How Computers Play Chess, Springer(1982)
- [12] Monty Newborn: Kasparov versus Deep Blue: computer chess comes of age, Springer(1997)
- [13] 松原仁: Deep Blue はなぜチェス名人に勝てたのか?そしてなぜ勝てなかったのか?, 情報処理学会誌, Vol.37, No.6, pp.536-542(1996)
- [14] 松原仁: なぜチェス名人はコンピュータチェスに負けなければならなかったのか?, 情報処理学会誌, Vol.38, No.8, pp.705-706(1997)
- [15] Chess.com “Komodo Overwhelms Erenburg In Odds Rematch - Chess.com”
<https://www.chess.com/news/komodo-overcomes-erenburg-in-2-pawn-odds-rematch-9996>
- [16] 村松正和: コンピュータ囲碁の現状, 情報処理学会誌, Vol.53, No.2, pp.133-138(2012)
- [17] Rémi Coulom: Efficient selectivity and backup operators in Monte-Carlo tree search, International Conference on Computers and Games, pp.72-83(2006)

- [18] ICGA “11th Computer Olympiad, Go (9x9) - Turin 2006 (ICGA Tournaments)”
<https://www.game-ai-forum.org/icga-tournaments/tournament.php?id=1>
- [19] David Silver and Aja Huang and Chris J. Maddison et al.: Mastering the game of Go with deep neural networks and tree search, Nature, Vol.529, pp.484-489(2016)
- [20] 大槻知史 “囲碁 AI “AlphaGo” はなぜ強いのか? ～ディープラーニング、モンテカルロ木探索、強化学習～”
<http://home.q00.itscom.net/otsuki/20160415AlphaGopublic.pdf>
- [21] 吉田裕之, 伊藤尚洋 “AlphaGo のしくみ”
<http://www.slideshare.net/yuk1yoshida/alphago-61311712>
- [22] 上海日報 “Lee Sedol expects "not easy" game with AlphaGo in 3rd Go match | Shanghai Daily”
http://www.shanghaidaily.com/article/article_xinhua.aspx?id=322918
- [23] 棋譜う | alpha の棋譜一覧
<http://www.kihuu.net/index.php?type=3&key=alpha>
- [24] WIRED.jp, “黒 37 手と白 78 手 : AlphaGo とイ・セドルが再定義した「未来」 WIRED.jp”
<http://wired.jp/special/2016/alphago-vs-sedol/>

- [25] David Ormerod, “Lee Sedol defeats AlphaGo in masterful comeback
- Game 4”
<https://gogameguru.com/lee-sedol-defeats-alphago-masterful-comeback-game-4/>
- [26] David Silver, Thomas Hubert, et, al.: Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm, ArXiv e-prints, arXiv:1712.01815(2017)
- [27] David Silver, Thomas Hubert, et, al.: A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, Science.
Vol. 362, Issue 6419, pp.1140-1144(2018)
- [28] 日本将棋連盟 “電王戦 | 棋戦 | 日本将棋連盟”
<https://www.shogi.or.jp/match/denou/>
- [29] “オセロプログラムと人間はどっちが強いのか？ ロジステロとの戦い”
<https://uguisu.skr.jp/othello/7-2.html>
- [30] “Yixin”
<https://www.aiexp.info/pages/yixin.html>
- [31] 三木理斗, 三輪誠, 近山隆: UCT 探索による不完全情報下の行動決定,
ゲームプログラミングワークショップ 2009 論文集, Vol.2009, No.12, pp.
43-
50(2009)

[32] Dwango Media Village “深層学習麻雀 AI「NAGA」 - Dwango Media Village(ドワンゴメディアヴィレッジ,dmv)”

https://dmv.nico/ja/articles/mahjong_ai_naga/

[33] “人狼知能プロジェクト”

<http://aiwolf.org/>

[34] 梶原健吾, 鳥海不二夫, 稲葉通将ら: 人狼知能大会における統計分析と SVM を用いた人狼推定を行うエージェントの設計,人工知能学会全国大会論文集, 第 30 回全国大会(2016)

4 章 本研究が対象とするゲーム

本研究では不完全情報ゲームの一種である「花札」を研究の対象とした。本章ではその花札の遊戯方法について述べる。

4.1 花札とは

花札とは、かるたの一種であり、古くから遊ばれてきた日本の伝統的なゲームである。その遊戯方法は地域によっても異なるが、大まかには「めくり系」と「カブ系」に分けられる。めくり系とは、プレイヤー間で札を取り合う体系の遊戯で、ここには「こいこい」「花合わせ」「六百間」などが存在する。一方のカブ系は配られた札の特定の数へ近づける遊戯で、「オイチョカブ」「京かぶ」「五枚株」などがある。他、韓国花札(花鬪)の系列に「ゴーストツップ(めくり系)」「ソッタ(カブ系)」がある。

以上のように花札と一口に言っても数多くの遊戯方法があるが、本研究では市販ゲームで最もよく採用される「こいこい」を選択した。

4.2 花札の札種類

図 4-1 に花札の札一覧を示す。図 4-1 の通り花札はトランプのようにスーツが別れている。トランプの場合は 13 枚×4 スートの 52 枚であるが、花札は 4 枚×12 ヶ月の 48 枚である。

スーツの呼称は図 4-1 の右より「松・梅・桜・藤・菖蒲・牡丹・萩・芒・菊・紅葉・柳・桐」で、それぞれ 1 月・2 月…12 月に対応する。めくり系の遊戯方法では、松の札と松の札、芒の札と芒の札といったように同じ月の札を合わせると札を取ることができる。この札を取り合うことが遊戯の基本である。

また、札にも種類があり「光札・種札・短冊札・カス札」と呼称する。「光札」は「松に鶴・桜に幕・芒に月・柳に小野道風・桐に鳳凰」の 5 枚である(図 4-2)。

「種札(タネ)」は「梅に鶯・藤に不如帰・菖蒲に八橋・牡丹に蝶・萩に猪・菊に盃・紅葉に鹿・柳に燕」のことを指す(図 4-3)。「短冊札(タン)」は札に短冊が描かれている札を指す(図 4-4)。以上に属さない札を全て「カス札(カス)」と呼ぶ。この札の種類分けはこいこいで役を作る際に必要となる。

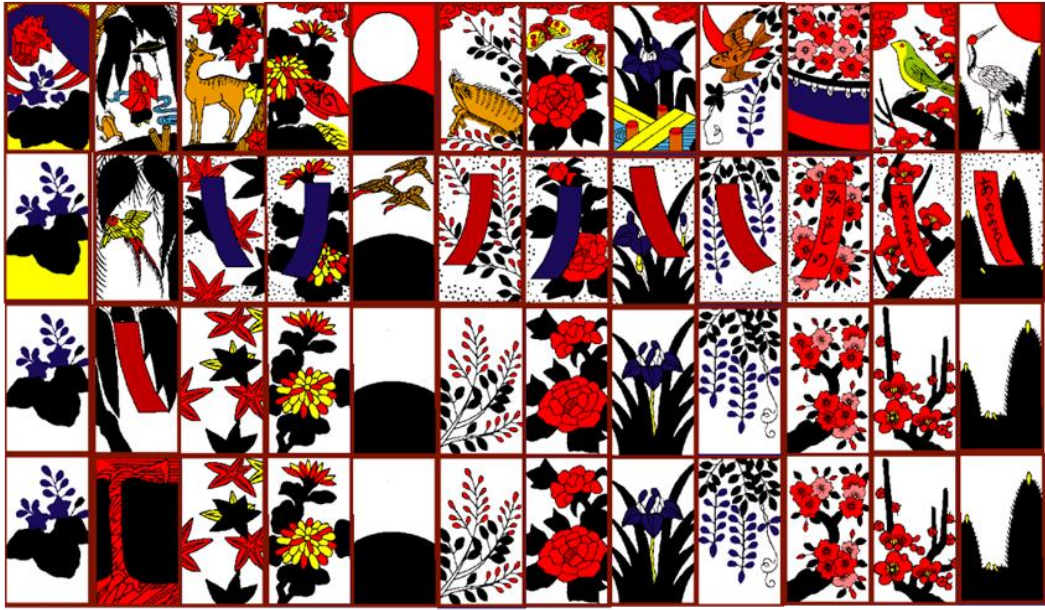


図 4-1 花札の札一覧



図 4-2 光札



図 4-3 タネ札



図 4-4 タン札

4.3 こいこいの遊戯方法

こいこいとは、2人のプレイヤーが親と子に分かれて札を取り合い、規定の役を作るゲームである。役は十数種類あり、これらは特定の札を集めることか、種類が同じ札を一定の枚数獲得することによって成立する。また役には点数が定められており、その単位は「文」である。本研究が採用した役については表 4-1 にまとめる。

ゲームの流れは以下の通りである。ここでは実物の札を用いて行う場合について述べる。

① 双方が 1 枚札を引き、親を決める。月が早い方の札を引いた者が初回の親となる。

② 親は双方の手札に 8 枚、場札に 8 枚を配る(図 4-5)。

この時一度に配らず相手に 4 枚、自分に 4 枚、場に 4 枚を二回繰り返す。

残りの札は山札として伏せておく。

③ 配り終わったら親から以下の行動を繰り返す。

(ア) 手札から 1 枚選び、場に出す。この時手札から出した札といずれかの場札の月が一致している札があるならば、その上に軽く打ち付ける。一致している札が複数あるならばどれか 1 枚を選んで打つ。その後、山札をめくり同様にし(図 4-6)、2 枚を組として取り札とする。取り札は種類別に分けて自分の前に置いておく(図 4-7)。一致している札がない場合はそのまま場に捨てる(図 4-8)。












(イ) この段階で役が完成しているときは「こい」と宣言し対局を続行するか、「上がり」と宣言し対局をやめるかの選択を行う。続行したなら、

新たな役ができるまで続け文数の上積みを目指す。やめたならば完成した役の文数を相手より受け取り，上がったプレイヤーが親になり②に戻る。役が完成していない場合は手番を交換し(ア)に戻る。

- ④ 双方が手札を出しきっても役が完成していないならば「親権」として親に 6 文の役を与えて②に戻る。この場合の次の親はそのままとする。

どちらかが上がるまでを 1 局とし，12 局終了時点での文数が多い方が勝ちとなる。この 12 局をひとまとまりにして 1 対局と呼称する。

表 4-1 花札の役一覧

役名	文数	説明	組み合わせ
五光	15	光札を 5 枚	
四光	10	光札のうち柳を除く 4 枚	
雨四光	8	光札のうち柳を含む 4 枚	 (1 例)
三光	6	光札のうち柳を除く 3 枚	 (1 例)
猪鹿蝶	6	萩に猪・紅葉に鹿・牡丹に蝶の 3 枚	
表菅原	6	松に鶴・梅に鶯・桜に幕の 3 枚	
赤短	6	松のタン・梅のタン・桜のタンの 3 枚	
青短	6	牡丹のタン・菊のタン・紅葉のタンの 3 枚	
七五三	6	萩に猪・菖蒲に八橋・桜に幕の 3 枚	
花見で一杯	5	桜に幕・菊に盃の 2 枚	
月見で一杯	5	芒に月・菊に盃の 2 枚	
タネ	1～	種札 9 枚のうち任意の 5 枚	(省略)
タン	1～	短冊札 10 枚のうち任意の 5 枚	(省略)
カス	1～	カス札 24 枚のうち任意の 10 枚	(省略)

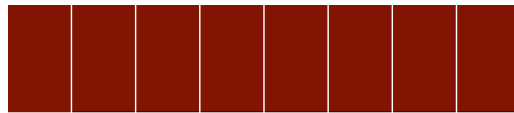


図 4-5 こいこい: 初期状態

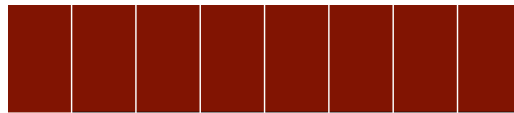


図 4-6 こいこい: 札取り

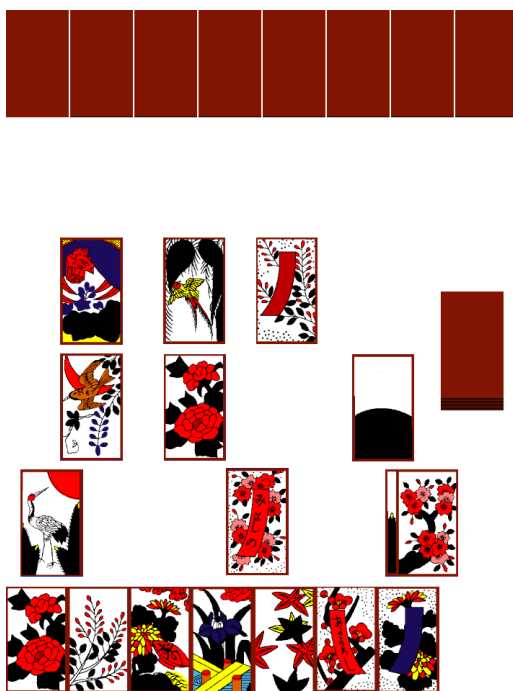


図 4-7 こいこい: 取り札の整理

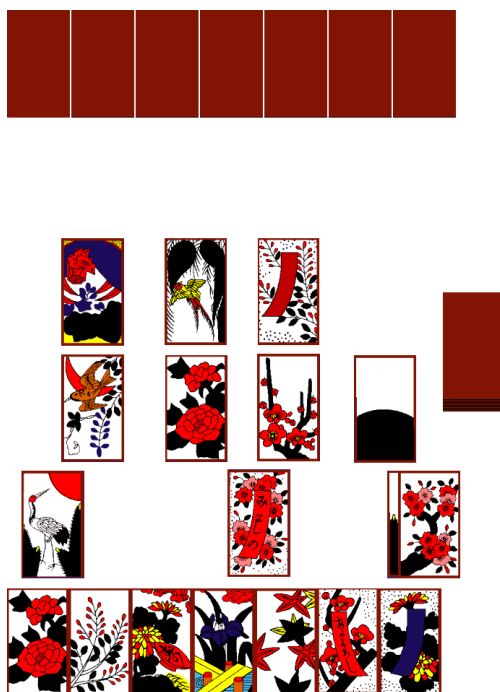


図 4-8 こいこい: 札が取れない場合

5 章 モンテカルロ木探索での強さの表現

本章では、ゲーム AI によく用いられるモンテカルロ木探索について述べ、モンテカルロ木探索を用いてゲーム AI を強くする研究を示す。

5.1 モンテカルロ法の初期: 原始モンテカルロ法

一般に「モンテカルロ法」と言った場合、多数の乱数を用いてシミュレーションや数値計算を行うことによって確率的に結果を求める手法である。これをゲームに適用するならば、ランダムにゲームを進行させ、結果を蓄積することで最善手を求める手法となる。具体的には、

- ① 局面を定める
- ② ゲームのルールに従って終局までランダムにプレイする(プレイアウトと呼称する)
- ③ 勝敗や得失点など、ゲームに使われる数値を記録する
- ④ ②, ③を 1 ループとして規定回数行う
- ⑤ 最も適した手(勝率の高い手であったり, 得点の高い手であったり)を選択するとなる。こういった手法で意思決定を行うものをモンテカルロ法, あるいは UCT と区別するために便宜上「原始モンテカルロ法」と呼んでいる[1]。本稿でも「原始モンテカルロ法」と呼ぶこととする。

原始モンテカルロ法は局面の良し悪しを評価する評価関数の作成が不要なのが利点である。評価関数がないため評価値の「算出」はできないが、プレイアウト結果から「推定」することが可能である。勝率を重視するようなゲームであれば

勝率を評価値とすれば良いし、得点を稼ぐゲームならば得失点を評価値とすれば良い。評価関数が不要であるという利点を受け、評価関数の作成が難しいとされているコンピュータ囲碁で注目された[2]。しかし欠点として、

- 乱数を用いるため不確実性が高く、相手のミスを期待した手選択をしてしまう
 - プレイアウト回数を大きくしても一定回数以上では実力が変わらない
 - 深い木に対しては最善手を返す保証がない
- という点がある[1]。

原始モンテカルロ法のモデルを図 5-1 に示す。ルートに現状の局面を置き、子ノードには自分の手番で取ることのできる手を置く。ゲーム木の作成が終了した後、全ての子ノードを均等に選択し、プレイアウトを行って局面の評価をする。プレイアウトの回数が増加するに従って評価値が収束していくので、取ることのできる手が良い手か悪い手かの判別ができる。しかし、均等に手を選ぶためどんなに悪手であっても試行されてしまい、非常に無駄が多い手法であると言える。

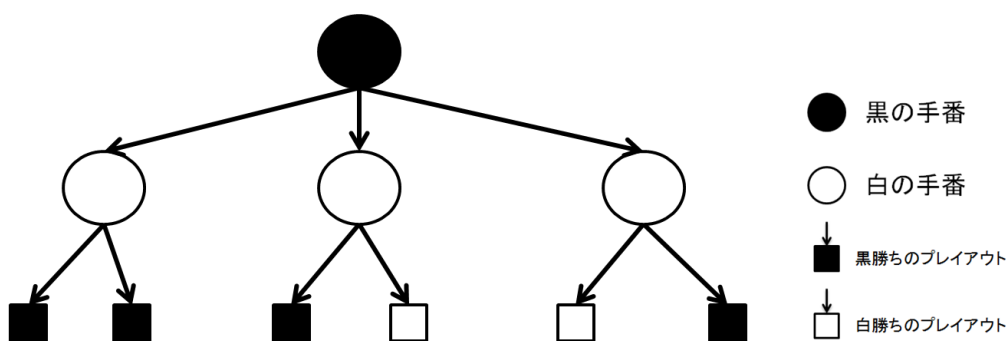


図 5-1 原始モンテカルロ法のモデル

5.2 原始モンテカルロ法の改良: モンテカルロ木探索

原始モンテカルロ法を改良したものがモンテカルロ木探索である。プレイアウトを行って評価をするのは変わらないが、図 5-2 のように有望だと思われる手を重点的に探索したり、図 5-3 のようにノードの選択回数が閾値を超えた場合は子ノードを展開したりするのが特徴である。子ノードが展開されるとプレイアウトの開始点が 1 段深くなり、探索木がどんどん成長して行く。有望な手を深く読んで手を判別するのである。有望であるかそうでないかは評価値を用いて判断する。仮に勝率を評価値とした場合は勝率が最も高い手を選択し、得失点を評価値とした場合は最も利得の高い手を選択する。この改良により悪手を選択する回数が減り、良い手を効率的に探索できるようになった。

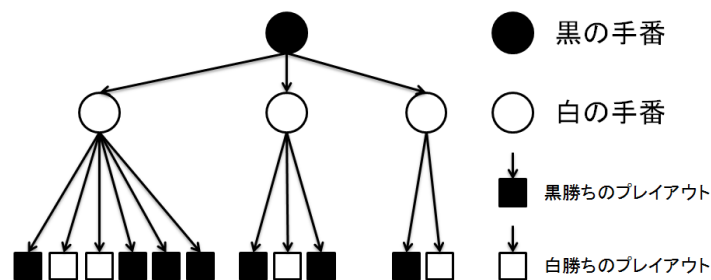


図 5-2 モンテカルロ木探索: 有望な手に多くのプレイアウト

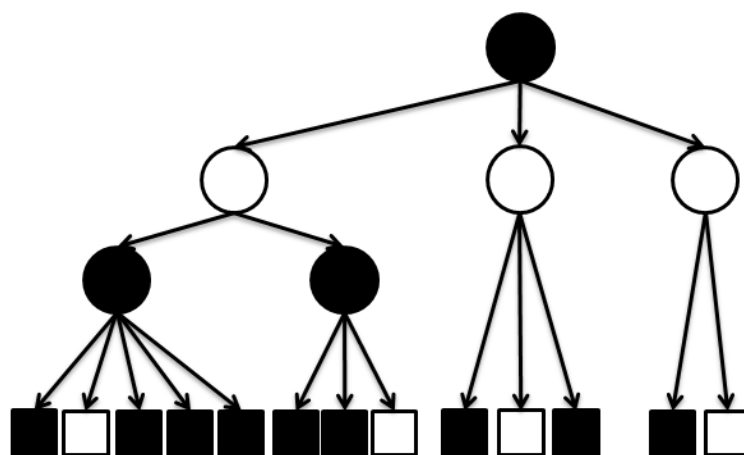


図 5-3 モンテカルロ木探索: 木の成長

5.3 木探索と UCB の組み合わせ: UCT の誕生

5.2 のモンテカルロ木探索と UCB を組み合わせたものが UCT である。UCT は 2006 年に Kocsis らによって提案された[3]。

5.3.1 UCB とは

UCB(Upper Confidence Bounds)は「多腕バンディット問題(Multi-armed Bandit)」を解くのに使用される戦略で、2002 年に Auer らによって発表された[4]。「多腕バンディット問題」は腕が複数あるスロットマシンのイメージで、各腕にコインを入れたら確率によって払い出しが受けられるようなものである。一度に投入できるコインは 1 つの腕のみであるので、どこにコインを投入すべきか考えるような問題となる。これをモデル化するならば、たくさんあるスロットマシンの中からどこにどれだけのコインを投入すれば儲けが出るか、となる図 5-4。これは「スロットマシンをするギャンブラーのモデル」と呼ばれる。

UCB は式(5-1)で計算される。式(5-1)中の \bar{X}_i は i 番目のマシンのその局面での報酬期待値、 n_i は i 番目のマシンのコイン投入数、 n は全てのマシンのコイン投入数の合計である。 c は定数で、この値が大きくなるとコイン投入数が少なく探索が進んでいない台を重視し、小さくなると報酬の高い台を重視するモデルであるが、よく用いられる値は $c=1$ である。式(5-1)により、単純に報酬期待値が高いマシンは選択されやすいが、期待値が低くともコイン投入数が少ないマシンは、第 2 項によって「ひょっとしたら優秀な台なのではないか」ということを考慮することができる。

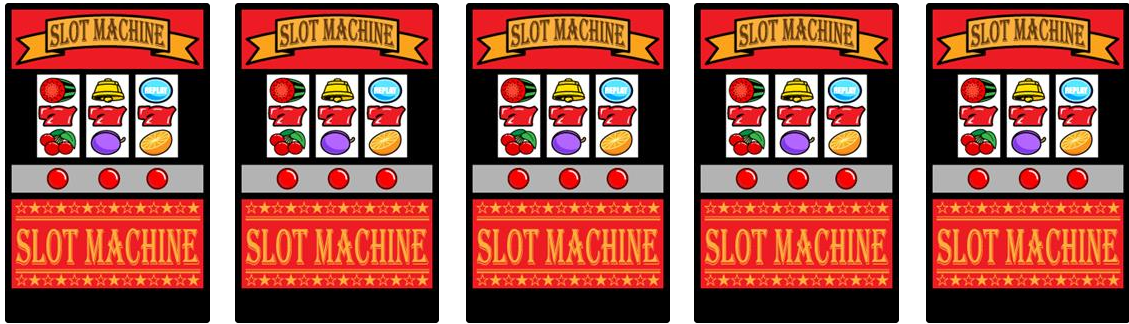


図 5-4 多腕バンディット問題のイメージ

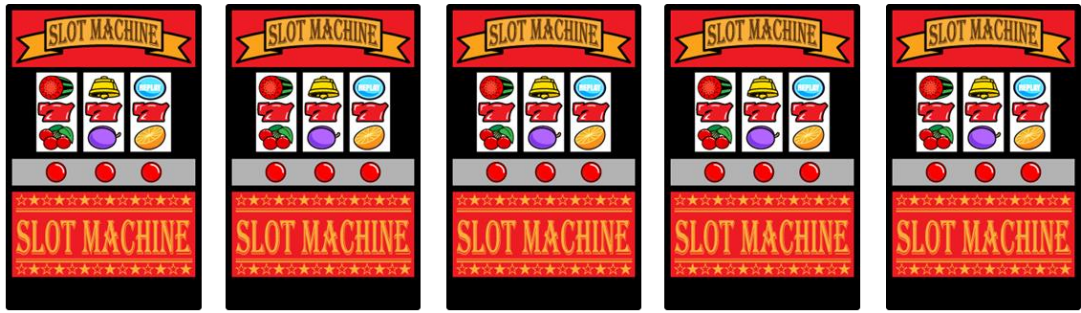
$$UCB(i) = \bar{X}_i + c \sqrt{\frac{2 \ln n}{n_i}} \quad (5 - 1)$$

5.4 UCB を利用した多腕バンディット問題の解法

ギャンブラーは次のような手順でコインを投入し、儲けを出そうと試みる。

- ① 全てのマシンにある程度の量のコインを等しく投入し、払い出しを記録する。
- ② 式(5-1)に従って UCB を計算する。式(5-1)中の \bar{X}_i は払い出しの平均値とすればよい。
- ③ その後は UCB が最も高いマシンを選択してコインを投入し、UCB を更新する。

具体例が以下の図である。ある時点での UCB が図 5-5 だとする。ギャンブラーは UCB が最も高いマシンをプレイする。その結果払い出しがあったりなかったりし、平均値の変化に伴って UCB が更新される(図 5-6)。仮に UCB が減少したとすると、他のマシンが最も高くなったのでギャンブラーは違うマシンをプレイする。これを繰り返すことによってマシンの選別が進み、良いマシンへ集中的にコインを投入することができる。



UCB: 1.142626

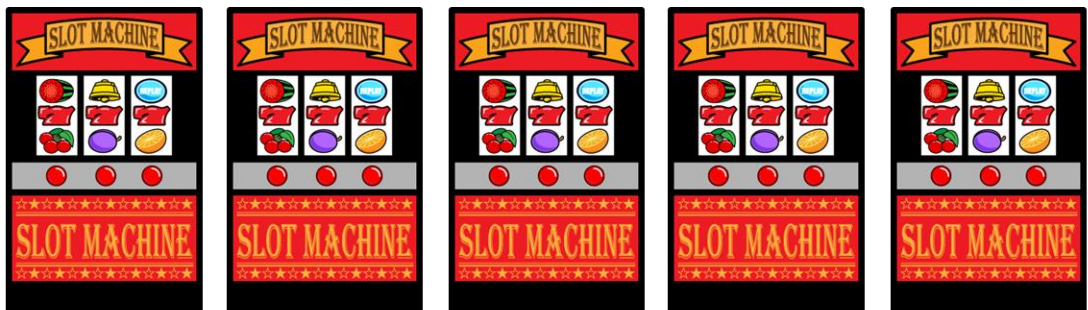
1.143526

1.137983

1.142368

1.146798

図 5-5 ある時点での UCB



UCB: 1.142626

1.143526

1.137983

1.142368

1.112368

図 5-6 UCB の更新

5.4.1 木探索へ応用

UCB をモンテカルロ木探索へ応用したものが UCT で、期待値の推定に UCB を適用するのである。UCT の概要図を図 5-7 に示す。図 5-7 における丸囲み数字が UCB である。

UCT はモンテカルロ木探索の各ノードを多腕バンディット問題であると見なし て UCB を計算して木探索を行う。木探索の方法は、

- ① 未探索のノードがあればそれを優先して選択, なければ UCB の値が最も高いノードを順に選択し, 選択した経路を保存する(図 5-8)。同時にノードの選択回数をインクリメントする。
- ② 木の末端まで到達したならばプレイアウトを行って勝負結果を記録する。ここでは利得を考慮することとする。
- ③ 利得をたどったノード全てに加え(図 5-9), ノードの UCB を更新する(図 5-10)。
- ④ ここまでで選択回数が閾値を超えたノードがあれば展開する。

このようにすることで、探索が進んでいないノードであっても式(5-1)の第二項によって考慮することができる。その結果、本質的には良い手であるが探索が進んでいないために見過ごされている手を選択することが可能となる。

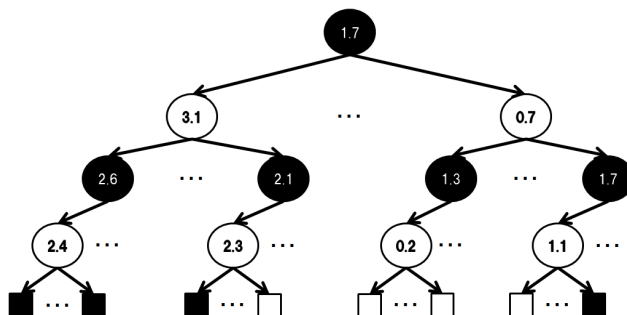


図 5-7 UCT 概要図

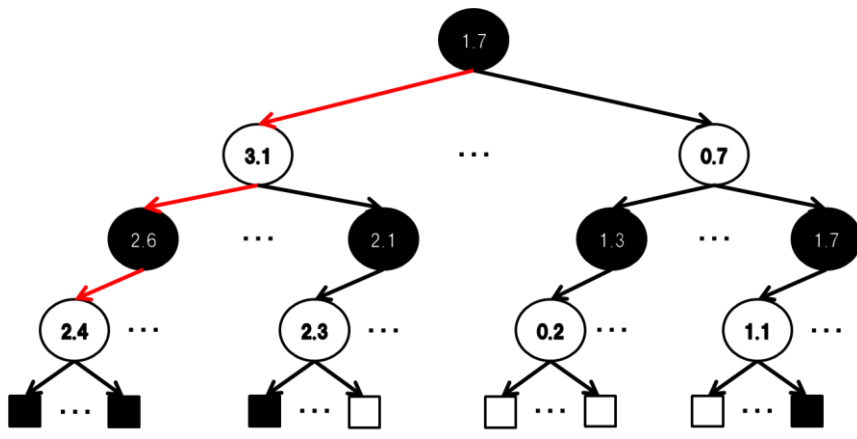


図 5-8 UCT: ルートを選択

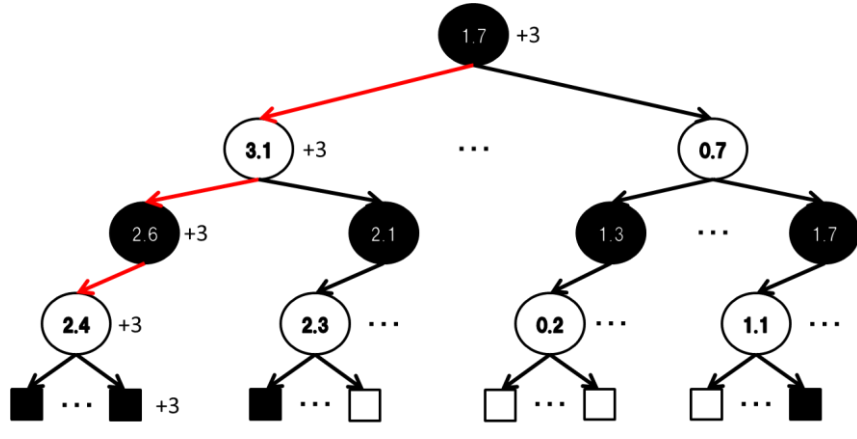


図 5-9 UCT: 利得を加算

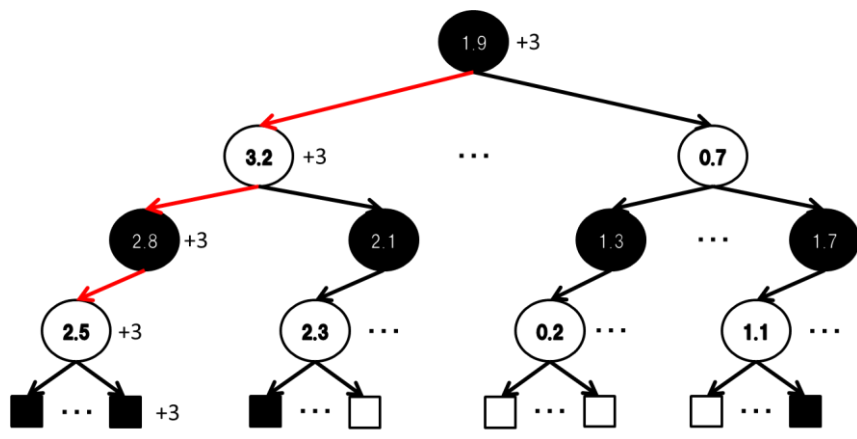


図 5-10 UCT: UCB を更新

5.5 UCT の適用

前述の UCT を利用して、花札の AI を強くする研究を行った。本研究では、自分の手番の時にどの手札を切るべきか選択する所に UCT を採用した。完全情報ゲームで用いられる Min-Max 法での局面の判断には、盤面の良し悪しを判断する評価関数の設計や、評価関数を作成するために対局の手順を記入した囲碁将棋における棋譜が必要となる。しかし、花札では棋譜を残すのは一般的ではなく、また研究が浅いためゲームの知識が集まっていないのが現状で評価関数の設計が困難である。一方、UCT はゲームの知識は必要とせず、シミュレーションの結果によって切る札の選択を行うので、評価関数の設計が困難である花札にも適用できると考える。

UCT 探索は以下の流れで行う。

- ① 相手の手札を推定する。
- ② UCT を一回行う。
- ③ ①, ②を規定回数繰り返す。
- ④ UCB が最も高い手を選択する。

この流れを図で表したものが以下の通りである。ここでは図 5-11 の局面を例にとる。まず相手の手札の推定を行う。相手がどのような札を所持しているかが不明であるため、見えている札から矛盾が出ないように、具体的には自分と相手の取り札と自分の手札を除いた札から相手の手札を推測する。アルゴリズム上では、一度相手の手札を山札に戻し、同じ枚数山札より引くことによって手札の無作為化を行っている(図 5-12)(当然だが UCT 探索を行うプレイヤーが相手の真の手札を知ることはできない)。続いて UCT のループを一回行う。本研究のゲーム木はル

ートに自分の手札の1枚を置き，その1段深いところに相手の手札を置く．木の末端ではプレイアウトが行われる．この木を手札の枚数分作成し，各手札を探索する(図 5-13)．なお本来はさらに1段深いところに自分の手札を置く必要があるが，現在では木を成長させずに1段読みのみでシミュレーションを行っている．すなわち，図 5-2 の木を図 5-3 のようにする手順は行わないものとする．また，本来ならばノードは局面である必要があるが，山札まで考慮した局面とすると局面数が増大するため，簡便な探索になるように手札をノードとした．

探索中のプレイアウトはどちらかが上がるか，双方が手札を全て出しきり「親権」によって終局となるまでとした．また 4.3 の通り，本来は役が完成しても更なる役を作れると判断すればゲームを続行することが可能であるが，本研究ではどちらかが役を完成させた時点でゲームを終了し，役代を受け取り次局に移ることとした．

①，②の手順を規定回数行った後，各手の UCB を比較して最も値の高い手を選択する(図 5-14)．この局面では「芒のタネ」の値が最も高いので，「芒のタネ」を選択する．以上が本研究における UCT 探索の流れである．

花札のゲームの目的は相手よりも多くの文数を獲得することである．UCT では式(5-1)の報酬期待値に勝率を用いるのが通常であるが，本研究では役が完成して獲得した文数を用いた．完成できる役が複数ある場合，勝率による判断ではいずれも同じ値を返すが，このようにすることで同じ勝利でもより報酬の大きな手が選択されるのではないかと考える．

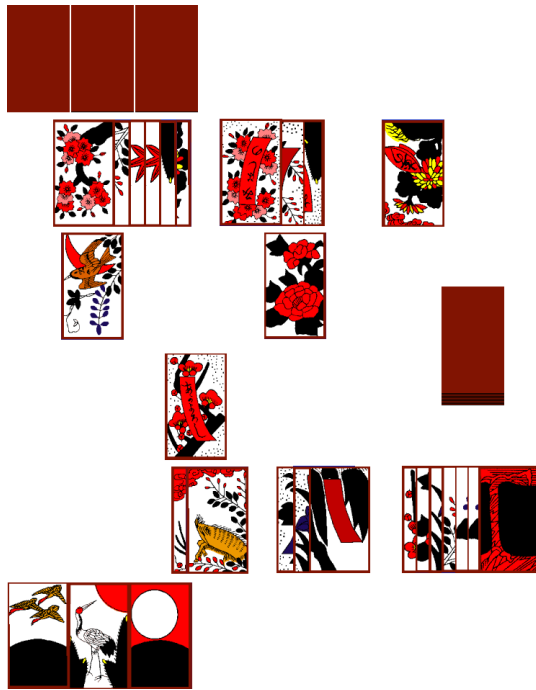


図 5-11 探索する局面

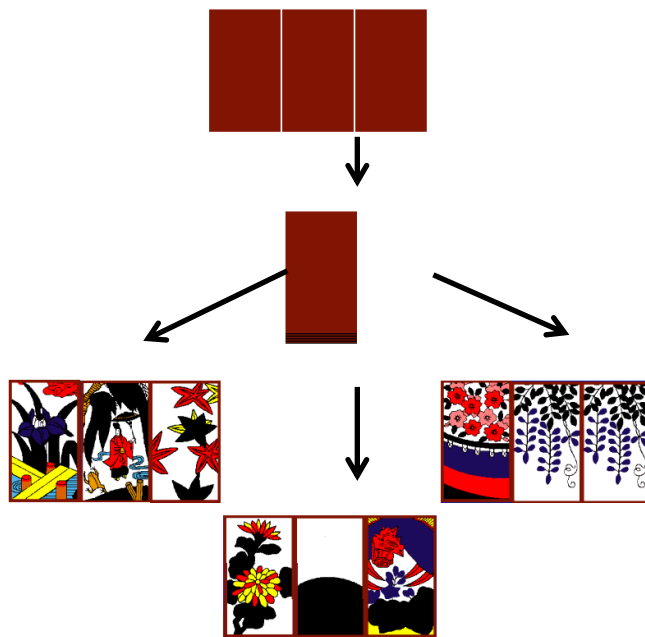


図 5-12 相手の手札の推測

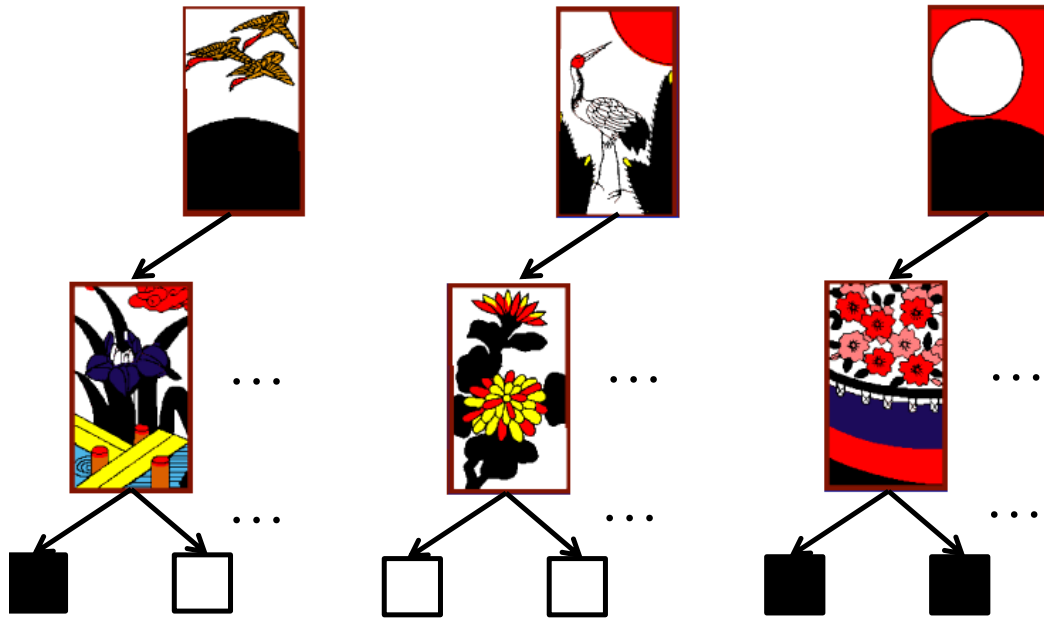


図 5-13 UCT 探索木

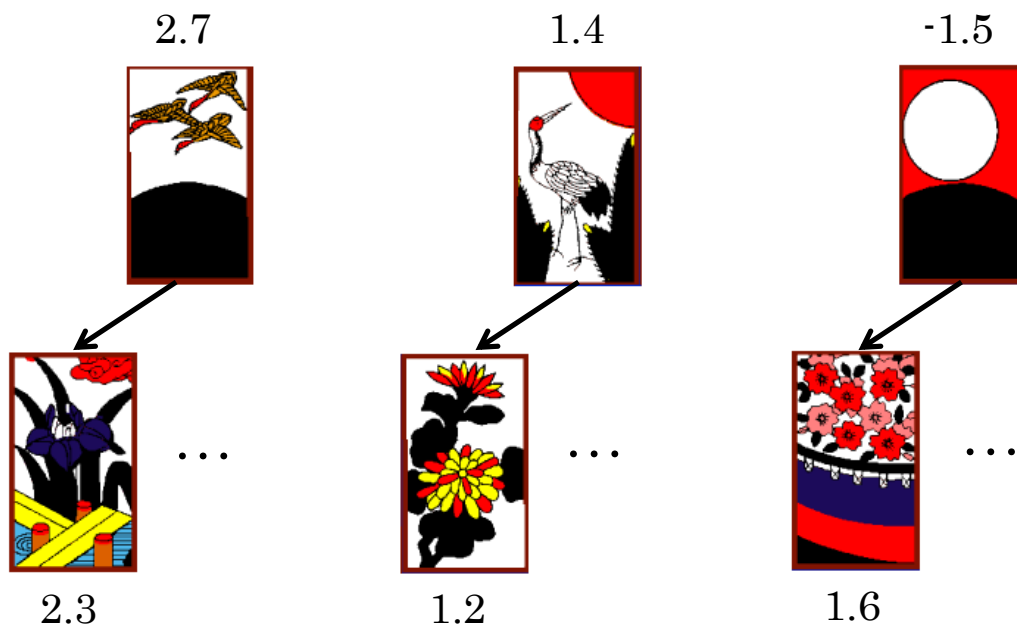


図 5-14 手札の決定

5.6 実験

手札の選択に UCT を採用したコンピュータプレイヤーを作成し、人間プレイヤーとの対局実験を行う。実験の設定は以下の通りである。

- 式(5-1)中の $c = 1$, \bar{X}_i : 獲得文数の平均
- UCT ループ回数: 2000
- 実験回数: 20 対局

実験は花札に慣れている人間プレイヤーで行い、正確な評価が行えるように努める。

20 対局行った獲得文数のグラフを図 5-15 に示す。横軸が対局回数、縦軸が UCT プレイヤーの獲得文数を表している。文数がプラスの場合は UCT プレイヤーの勝利、マイナスの場合は人間プレイヤーの勝利と定めると、図 5-15 からは UCT プレイヤーの 14 勝という結果が見える。またこの実験において、1 対局の平均獲得文数は UCT プレイヤーの 7.00 文勝ちであった。特に、19 回目の対局では 52 文の勝ちを収めており、強い実力を見せている。したがって、UCT プレイヤーが人間プレイヤーに勝ち越していることから、UCT が花札にも有効である可能性があることが得られた。



図 5-15 UCT プレイヤーの獲得文数

実験の中で、特徴的だと思われる UCT プレイヤの打ち方がいくつか確認できたためここに掲載する。図 5-16 は対局中のある局面である。この状況は、UCT プレイヤが花見で一杯と月見で一杯を同時に完成させようとしていて、手札には役の完成に必要な菊のタネの札を持っている。さらに場札には菊のカスがあり、すぐにでも役の完成が見込まれる状況となる。この時の UCB が表 5-1 である。当然のことながら、菊に盃の札の UCB が最も高くなっている。

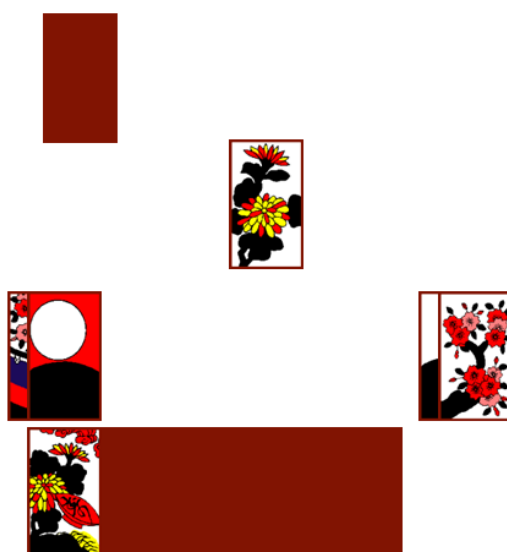


図 5-16 ある局の場札と UCT プレイヤの手札

表 5-1 図 5-16 の局面の UCB

オーダー	選択	UCB
1(最善手)	菊のタネ	12.13
2	Other	2.42
...

図 5-16 の局を開始した時の状況が図 5-17 である。手札や場札に月見・花見を構成する札が揃っている。この状況下において、通常ならば桜の光札を優先的に取得し、得点をより多く獲得しようと試みる場面であるが、探索が終わった時の UCB が

表 5-2 となった。探索の結果、セオリー通りに桜のカスを最善手としていて、正しい探索が行えているといえる。



図 5-17 図 5-16 の局の開始時点

表 5-2 図 5-17 の局面の UCB

オーダー	選択	UCB
1(最善手)	桜のカス	7.346
2	芒の月光札	4.746
3	Other	3.196
...

一方，獲得文数の平均を報酬期待値としたために生じる問題も確認されており，
図 5-18 はその典型的な状況である．図 5-18 は UCT プレイヤが七五三を狙える
状況だが，場には桜の光札の札が存在しない．ここで，通常ならば手札の桜のカ
スを最後まで温存して札が場に出るのを待つのがセオリーである．しかし，探索
の結果，他に切るべき札がない場合はこのキーカードを切ってしまうプレイング
がしばしば見られる．これでは有利な状況が一転して不利になるため，この部分
を改善することが必要である．その例としては，キーカードを取得できない時は
探索の対象から一時的に除外することや，手札をロックして切ることをできなく
させることが考えられる．

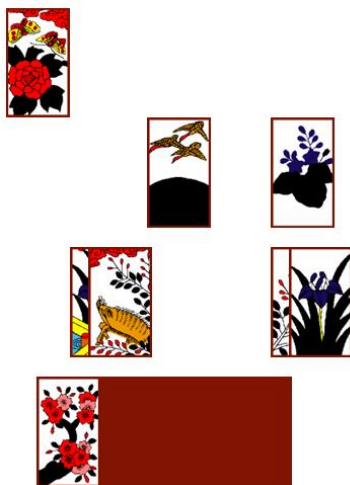


図 5-18 好ましくない札を選択してしまう局面

5.7 まとめ

花札の AI を強化する研究として，囲碁や将棋の分野で顕著な実力を示している
UCT による意思決定を行うコンピュータプレイヤを作成し，人間プレイヤとの対

局実験を行った。20回の対局実験の結果、コンピュータプレイヤーが14勝を挙げ、1対局平均7.00文を獲得できることより、花札にもUCTが適用できる可能性があることが得られた。またUCTの挙動を解析したところ、セオリー通りに札を選択することができるが、一部の選択が不自然であることが判明した。従って、より強くするためには選択の不自然さを排除することが必要である。

参考文献

- [1] 美添一樹: モンテカルロ木探索-コンピュータ囲碁に革命を起こした新手法, 情報処理, Vol.49, pp.686-693(2008)
- [2] Bernd Brügmann: Monte Carlo Go, White paper(1993)
- [3] Levente Kocsis and Csaba Szepesvári: Bandit based Monte-Carlo Planning, ECML'06 Proceedings of the 17th European conference on Machine Learning, pp.282-293(2006)
- [4] Peter Auer and Nicolo Cesa-Bianchi and Paul Fischer : Finite-time Analysis of the Multiarmed Bandit Problem, Machine Learning, Vol.47, pp.235-256 (2002)

6 章 ゲーム AI と面白さ

6.1 近年の動向

近年では、ひたすら強さを追い求めるのではなく、より人間プレイヤーを楽しませようとするゲーム AI の研究がなされている。これに関して池田は、「囲碁の初級者が上級者との対戦を行うとき、上級者が勝ちすぎてしまう結果として、初級者の上達途中にある段階で意欲を失ってしまう。これに対し、教えるプロは手加減を行い、楽しませる局面を導く技術を持つ。この技術を AI にも持たせたいというのが次の大きな目標である」と語っている [1]。このように、人間を楽しませようとするゲームをプレイする AI が作成できないかを研究する事例があり、AI を強化する研究とはまた異なった流れができています。本章では、これらの研究のうちいくつかを述べる。

6.2 面白さを表現する研究の例

6.2.1 サッカーゲームにおける NPC の行動検討

高橋らは、サッカーゲームの NPC が人間プレイヤーを楽しませることができる行動は何かを検討している [2]。この中では、高橋らが定めた面白さを与えられると思われる行動をとる NPC を 4 チーム分作成し、人間プレイヤーが操作する一人と共闘させて行動を評価する手法で実験が行われた。作成したチームは、

- T1: 人間プレイヤーの位置を優先的に観測する行動を取る NPC
- T2: ボール所持時、人間プレイヤーへ無条件にパスを送る行動を取る NPC

- T3: ボール所持時、人間プレイヤーへのパスがある程度決まりそうなら優先的にする行動を取る NPC

- T4: 消極的なフォーメーションを取る行動を取る NPC

である。いくつかの実験の結果、「楽しさが $T2 < T1$, $T4 < T1$, $T2 < T3$, $T4 < T3$, $T3 < T4$ となりやすい傾向にある」とし、「人間が NPC にある程度知的な行為を求めていることがわかる」と分析している。今後、より効果的な行動の調査や人間プレイヤーへ行動をどう認識させるかを検討したいとしている。

6.2.2 接待将棋・接待碁

仲道らはプレイヤーの棋力に合わせる接待将棋 AI を作成し、インターネット対局システムに投稿した[3]。これは探索で得られた局面の評価値を用いて、指し手ごとに評価値を 0 に近づけるような手を打つものである。この AI と人間プレイヤーの対局実験を実施し、プレイヤーへアンケートを取り AI を評価している。その結果、AI が強すぎる・弱すぎると評価したプレイヤーは面白くないと評しており、AI が適度に強いとしたプレイヤーが AI を面白いと回答している。また不自然な選択は面白さを損なっているとし、人間プレイヤーと AI との棋力差により不自然さを感じると述べている。

池田らは囲碁を対象とした接待システムの構築を検討している[4]。この中では接待碁に必要な要素技術をまとめ、それらがどのように作用するかを記述している。その後モンテカルロ法を用いた面白さの演出を目指し、改良すべき点を挙げている。

両ゲームに共通するのは、面白さを与えるためには不自然な着手をなくすことが求められると述べている。またコンピュータプレイヤーは強すぎても弱すぎてもよくないもので、人間プレイヤーの実力に合った選択をする必要があるとしている。

6.2.3 対戦格闘ゲームの AI がもたらす面白さ

石原ら是对戦格闘ゲームの AI や操作法がプレイヤーの感じる面白さに与える影響について分析している [5]。石原らは UCT とルーレット選択、およびルールベースの AI を構築、UCT の評価関数を変更することで人間プレイヤーに合わせた強さを表現し、AI がプレイヤーに与える影響を分析した。プレイヤーに対するアンケート調査により、やはり強すぎる AI は面白くないものであるとの結果を得ている。また不自然な動きは面白さを損なうとしたものの、それよりも互角の勝負をするほうが面白いと感じられるのではないかと結論づけている。ただし、行動の不自然さを取り除けばより良くなるとも述べている。

6.3 総括

いずれの文献においても、以下の事項が必要だと述べている。

- 強すぎても弱すぎてもよくない
- 選択が不自然であると面白さを損なう

従って同様に面白い AI を構築するならば、この事項に気を付けることが重要であろう。

参考文献

- [1] 池田心: 楽しませる囲碁・将棋プログラミング, オペレーションズ・リサーチ: 経営の科学, Vol.58, No.3, pp.167-173(2013)
- [2] 高橋克樹, 西野順二: サッカーゲームにおけるエンターテイナー NPC の行動, 第 33 回ファジィシステムシンポジウム 講演論文集, pp.575-578(2017)
- [3] 仲道隆史, 伊藤毅志: 人を楽しませる接待将棋システム, 人工知能学会全国大会論文集, 第 28 回全国大会(2014)
- [4] 池田心, Simon Viennot: モンテカルロ碁における多様な戦略の演出と形勢の制御～接待碁 AI に向けて, ゲームプログラミングワークショップ 2012 論文集, pp.47-54, 第 17 回ゲームプログラミングワークショップ(2012)
- [5] 石原誠, 宮崎泰地ら: 対戦格闘ゲームにおけるゲーム AI や操作法の違いがプレイヤーの感じる面白さに与える影響の分析, 情報処理学会論文誌, Vol.57, No.11, pp.2414-2425(2016)

7 章 面白さを表現するためのアプローチ

本章では人間プレイヤーに面白さを与えるコンピュータプレイヤー作成のために実施した研究について報告する。

7.1 研究背景

6 章で述べた通り，ゲーム AI を面白くすることも現在よく行われている．確かにゲーム AI の研究の本流は強くすることなのだが，対局していて感じられるのは「強すぎると面白くない」ということである．あまりにもコンピュータプレイヤーが強くて勝てないとなると，プレイヤーは退屈に感じるものであるというのが対局実験により判明した．そこで，面白さを与えるような花札のコンピュータプレイヤーは作成できるのかと考え，この研究を行うものである．

7.2 UCT プレイヤを改変した面白さの表現

第一に行った実験は、ゲーム AI を強化する目的で構築した UCT プレイヤを改変して面白さを表現するものである。UCT はある程度正確な探索を行える手法であり、独自に定めた面白さの定義に従って花札を打てば面白くなるのではないかということから研究を行った。

7.2.1 面白さの定義

本研究では面白い対局の定義を「得点のやり取りが大きく起こり、大勝ちせずに勝負感が演出できる対局」と定めた。花札で相手に勝利するためにやることは得点を稼ぐことである。そこでプレイヤは 1 文でも多く得られる札を選択し、相手が有利になるような札を切らないようにする。この状況で楽しむためには、小さい得点のやり取りではなく大きな得点を上げ、勝負をしているよう演出したらどうかと考えた。ただし、ひたすらに得点を稼ぐのでは強いプレイヤと変わらないので「大勝ちせずに」という制限を加える。よって、時にコンピュータプレイヤは自身が負けるような手を選択し、相手に勝たせるよう仕向けるようにするのが必要である。これにより対局の中で「得点を取った取られた」を繰り返し、最終的には小勝ちか小負けに留めるようなコンピュータプレイヤは面白いのではないかとした。

コンピュータプレイヤがこの定義に従って手を選択する時に考慮すべき事項は次の 2 点である。

1. 各局でやり取りされる得点を大きくする
2. 対局終了時の得点を ± 0 付近に調整する

1.の事項は「つまらない対局を避ける」ことを表す。得点のやり取りが1点や2点の繰り返しでは小刻みな得点のやり取りでしかなく、退屈な対局になりがちである。これを避けるため、コンピュータプレイヤーが得る、あるいは失う得点を大きくし、派手な対局を演出したらよいのではと考えた。実際の対局でも、得点のやり取りが小さい局(俗に「小場」という)と大きい局(「荒場」という)とでは、荒場の多い方が爽快感を感じられ、面白い印象になりがちである。そのことから、1.の事項を定めたものである。2.の事項は「対局の印象をよくする」ことを表す。前述の通り、コンピュータプレイヤーが強すぎると人間プレイヤーがつまらないと感じやすくなり適したものではない。一方で弱すぎても、歯ごたえがないと思われてしまい良いプレイヤーとは言えない。従って、大勝ちも大負けもしないプレイヤーが必要であるから、2.の事項を定めたものである。

7.2.2 UCT の変形

以上の定義を満たすようなコンピュータプレイヤー(以下「演出プレイヤー」と呼ぶ)を、UCTを変形することにより実現できないか考えた。具体的な変形方法は、式(5-1)で計算したUCBの利用方法を変えるものである。式(5-1)により計算されるUCB値は、本研究で利用した設定においてはおよそその手札の期待値に等しくなる。従って、この手札を選択したらどのくらいの文数を得られるか、あるいは失うかをある程度計算することができる。そこで、演出プレイヤーの各手番にて選択する手札を次のように決定する

$$Select\ hand = \begin{cases} \max_i(-UCB(i)) & \text{won in last time} \\ \max_i UCB(i) & \text{otherwise} \end{cases} \quad (7-1)$$

式(7-1)において、 i は手札の番号である。条件判定には、前の局でコンピュータプレイヤーが勝利したかどうかを使用する。例えば、演出プレイヤーが前の局で得点を得たとする。その場合には、測定した UCB が最も小さい札を選択し、人間プレイヤーの得点が最大になるような札を切ることとなる。このようにして、7.2.1で述べた事項の内 1.を達成しようとした。対局の中で、UCB が最大の札を切る局と最小の札を切る局とを交互に出現させ、各局でやり取りされる得点を大きくさせることを目指したのである。

この変形が機能するかの実験を、グリーディ方策に従って意思決定を行うコンピュータプレイヤー(以下「グリーディプレイヤー」と呼ぶ)を用いて実施した。グリーディ方策とは、取得できる札があるなら取り、ひたすら役を完成させることを目指すものである。このグリーディプレイヤーと演出プレイヤーとを対局させ、正しく得点を操作できるかを検証する。収集するデータは両プレイヤーが得た得点で、仮に演出プレイヤーが 6 文を得たならば+6 を記録し、グリーディプレイヤーが 5 文を得たときには+5 を記録する。得た得点を積み上げることにより、1 対局の中でどの程度大きな得点のやり取りができるかを観測する。

図 7-1 に 12 対局行ったときの得点の記録を示す。横軸が対局回数、縦軸がやり取りされた得点の総数である。それぞれの対局において、最終的に 60 文を超えるやり取りを達成できていることがわかる。また 90 文近くのやり取りがなされている対局も存在している。90 文のやり取りということは、1 局あたり 7 文前後の得点ということになり、毎局 6 文の役を完成させていることに相当する。6 文の役は特定の札を集めなければならず、1 文の得点で終わることもあると考える

と、毎局 6 文の役を作ることは難易度が高いものである。これを考慮すると得点の調整に成功していると言えるのではないか。

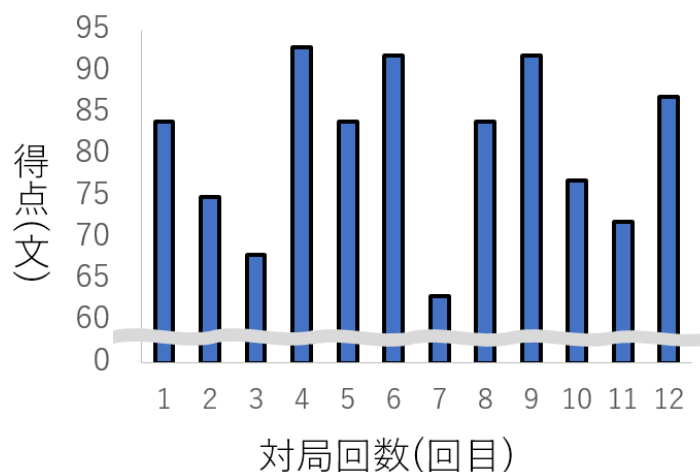


図 7-1 対局でやり取りされた得点

典型的と思われる 1 つの対局で得られた結果を図 7-2 に示す。横軸が局回数、縦軸が演出プレイヤーの得た得点である。得点がマイナスの場合は演出プレイヤーが失点したことを意味する。この対局ではやり取りされた得点の合計が 68 文であり、1 局あたり 5.67 文がやり取りされている結果となった。やり取りされた得点が 70 文付近となったことから、やり取りされる得点を大きくする目標は満たしている。また対局の中で勝つ局と負ける局をほぼ交互に繰り返しており、勝ちっ放しにはなっていない。従って UCT により得点の調整が行えていることがこの図より分かる。

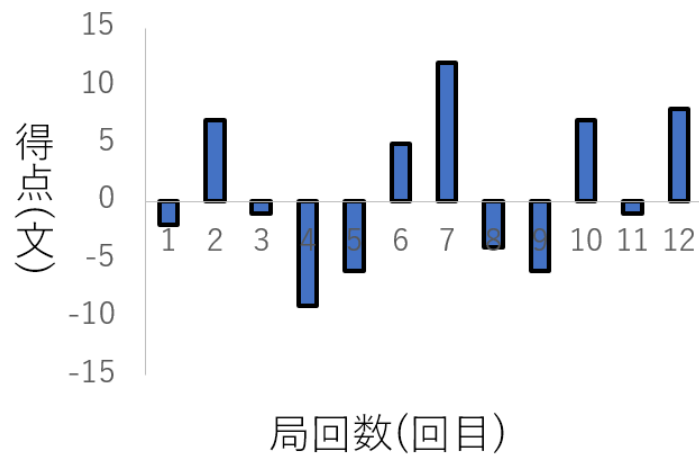


図 7-2 ある局の得点推移

しかし、演出プレイヤーが負ける番になったときには問題が生じる。現在の設定では UCB が最小の札を選択するようになっているが、ここに問題点が存在している。図 7-3 はある局で観測された演出プレイヤーが選択した札の UCB の変化である。図ではターンを減るたびに UCB が減少していることがわかる。演出プレイヤーはとにかく負けようとするため、式(7-1)より UCB が最も低い札を選択する。このため図 7-3 のように UCB が終始マイナスになるのだが、これは相手に得点をしてもらおうという札を毎ターン選択していることを意味している。つまり、演出プレイヤーの切る札があからさまであるということであり、仮にこの演出プレイヤーを人間プレイヤーと対局させたとき、あからさまな切り札をしてくるプレイヤーに面白さを感じるだろうか。予想ではあるが、あからさまで狙いがわかってしまう札を切ってくるプレイヤーには面白さを感じられないと思われる。よってこのままでは面白さを演出するプレイヤーにはならない。そこで、UCT により得点の調整が可能ではないかという結果を得たので、さらなる変形により面白さを演出できないかを検討した。



図 7-3 ある局の UCB の変化

7.2.3 UCT のさらなる改変

7.2.1 にあるコンピュータプレイヤーが考慮すべき事項を満たすように，UCB の利用方法をさらに変更し，演出プレイヤーの手札を式(7-2)により決定する．

$$Select\ hand = \begin{cases} \min_i |point + UCB(i)| & point > 0 \\ \max_i UCB(i) & otherwise \end{cases} \quad (7-2)$$

式(7-2)において，*point*は演出プレイヤーの得点，*i*は手札の番号である．式(7-1)と同様，演出プレイヤーは相手プレイヤーに得点させようとする札を切るモデルであるが，演出プレイヤーがわざと失点しようとするタイミングが異なっている．式(7-1)では演出プレイヤーが得点を得た局の次の局だけであり，演出プレイヤーが大きな得点を得ている場合，次の局で相手プレイヤーが1文しか得られなかったとしても，その次の局では演出プレイヤーは得点を得ようとする．したがって点差が離れてしまうケースがしばしば見られ，結果的に強いプレイヤーとなってしまうこともあった．式(7-2)による選択では，演出プレイヤーの得点がプラスの時に失点しよう

とするようにし、得点の乖離が起こらないように変更した。また選択する札は演出プレイヤーが得ている得点に近い札を選択するものとし、これによって最終的な得点を±0へ調整できないかを考えた。一例として、演出プレイヤーが7文のプラスである場合を挙げる。演出プレイヤーの得点がプラスであることから、式(7-2)によってUCBが-7に近い札を毎ターン選択していくこととなる。そして次の局でも演出プレイヤーの得点がプラスならば、再びUCBが得点に近い札を選択する。これを繰り返し、最終的な得点を±0付近に調整して対局を終えるようにする。

さらに変形した演出プレイヤーを利用して、人間プレイヤーとの対局実験を実施した。実験では得点の推移を計測し、対局終了時に人間プレイヤーへ聞き取り調査を行う。内容は対局の印象で、演出プレイヤーの良い点と悪い点について尋ねた。

図 7-4 は実験で観測されたある局の得点推移である。横軸が局回数、縦軸が演出プレイヤーの得た得点となっている。この対局は演出プレイヤーの3文勝ちという結果であった。3文勝ちは±0にある程度近づいており、考慮すべき事項の2.を満たしていると言える。しかし、考慮すべき事項の1.については満たしているとは言い難い。グリーディプレイヤーと人間プレイヤーという差のため、やり取りされた得点が比較的小さくなっており、大きくするという事項は満たしていないと考える。人間プレイヤーが相手でも10文程度のやり取りができるよう改善しなければならぬ。

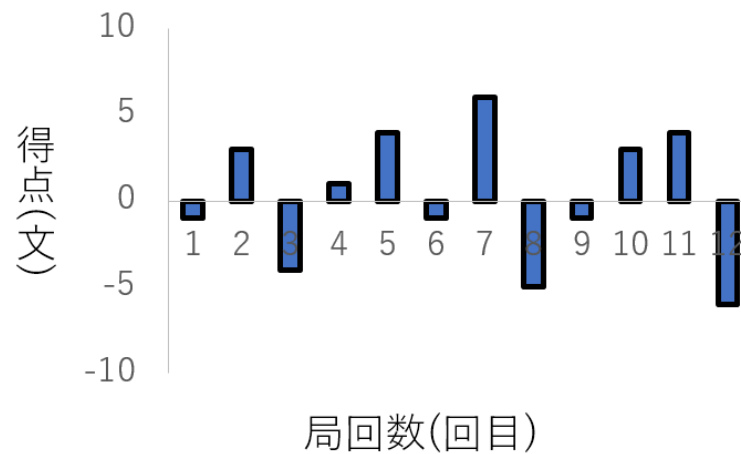


図 7-4 ある局の得点推移

また対局した人間プレイヤーのコメントは表 7-1 の通りである．コメントを分析すると次のことが言える．

- 勝てないほど強くはない
- 楽しませようというの見える
- 弱すぎるという人もいる
- 札の選択が不自然である

これより，一つわかることは人間プレイヤーに対しても勝ちっぱなしになっておらず，適度な手加減に成功していることである．また楽しませようとしているのは人間プレイヤーに伝わっているのがうかがえる．一方で，手を抜きすぎで弱すぎるというコメントもあり，さらには札の選択が不自然であることが見透かされている．つまりは札の選択があからさまであり，この不自然さを取り除かなければ面白さを与えるプレイヤーの構築がなされないということである．

表 7-1 対局者のコメント

対局者名	良かった点	改善点
A	<ul style="list-style-type: none"> ● 強すぎることはなく，なかなか勝つことができなかった。 ● 一方的な試合じゃなく，ある程度対局を楽しめた。 	<ul style="list-style-type: none"> ● コンピュータプレイヤーの札の選択が不自然である。こちらを勝たそうとしているのがわかる。
B	<ul style="list-style-type: none"> ● 対局を楽しませようとするのは見られた 	<ul style="list-style-type: none"> ● 強さが一定ではない。弱すぎるように感じた。

7.3 花札のプレイヤーへの面白さのアンケート

7.2の結果から、UCTでの得点調整には成功したものの、選択に不自然さがあることや弱すぎても良くないといったものが浮かび上がった。そして「面白さの定義」に関する議論もなされていなかったことより正確な「面白さの定義」の必要性が増すこととなった。この研究を行うにあたってはより正確な定義付けを行う必要がある。そこで本研究では花札における面白さを正確に定義すべく、第二の研究として「花札のプレイヤーが考える面白さは何か」の調査を行った。具体的には花札をプレイする人にアンケート調査を行い、その人の考える面白さとはなにかを調べるものである。その後結果を分析し、正確な定義づけを目指す。

7.3.1 調査方法と調査内容

調査はインターネット上のコミュニティサイトで参加者を募り、調査項目をメール本文で送付することにより実施した。回収はメールを返信してもらうこととした。実施期間は2018年3月16日から4月2日まで、調査に同意したのは76人、うち期間内に返送があったのは61人で回答率は80.3%であった。調査内容は以下の通りである。

① 調査対象者の特性

(ア) 年齢

(イ) 性別

(ウ) 花札をやっている年数

具体的な年数を記載してもらった

(エ) 花札を行う頻度

大まかな回数を記載してもらった

② 面白さに関して

(ア) あなたの考える面白さとは (自由記述・複数回答)

(イ) あなたの考える面白くないさとは (自由記述・複数回答)

(ウ) その他花札に対する考え (自由記述・複数回答)

これらの質問に答えてもらうこととした。

7.3.2 調査結果 1: 調査対象者の特性

(1)年齢および性別

まず性別であるが、男性 61 人で 100%であった。また年齢は幅広く、10代から50代までであり、平均は 31.9 歳であった。年齢の分布は図 7-5 の通りである。

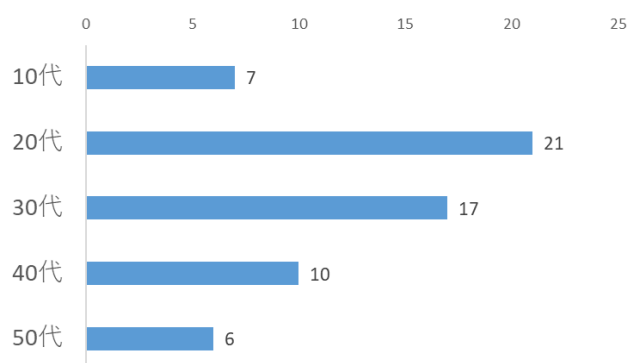


図 7-5 プレイヤの年齢(N=61)

(2) 花札をやっている年数

花札の経験を聞いた質問であるが、長い者で15年、短い者で1年半であった。

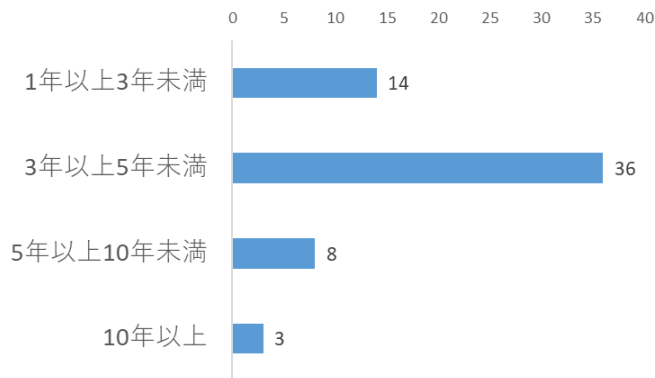


図 7-6 花札をやっている年数(N=61)

(3)花札を行う頻度

最も多かったのが週1回程度で、月1回程度、週3~4回、が続いた。その他、定期的ではないが暇になったらという回答や正月にのみ行うという回答もあった。なお、この回数は実際の札を用いて行う物のほかパソコン・スマートフォンその他電子機器上で行う物も含む。

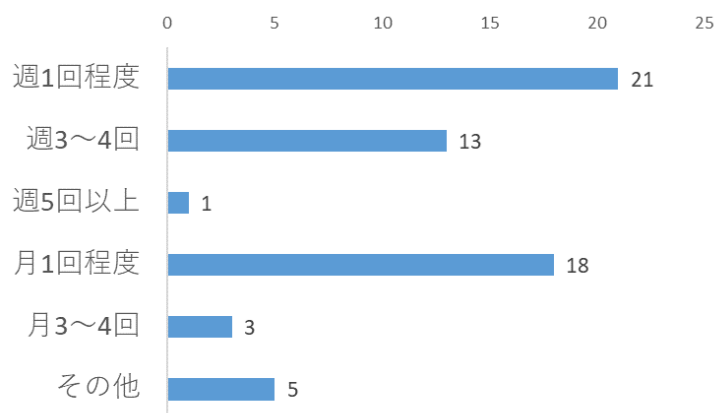


図 7-7 花札の面白いところ(N=61)

7.3.3 調査結果 2: 回答者が考える面白さ

(1)あなたの考える面白さとは

「あなたの考える面白さとは」という質問に自由記述で回答を求めた。整理した回答を以下に記述する。なお、複数回答したものはそれぞれについてカウントする（以下同様）。

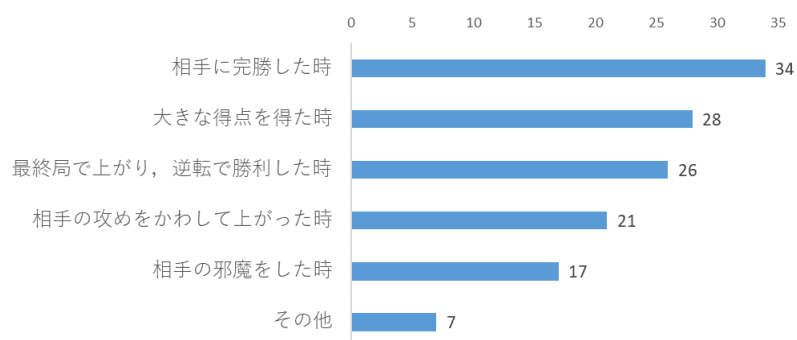


図 7-8 花札の面白くないこと(N=61)

- 相手に完勝した時
- 大きな得点を得た時
- 最終局で上がり、逆転で勝利した時

これらの項目が上位であった。その他にあたるものは「最後の1枚の手札で役が出来た時」や「最終局で相手を抑えた時」などがあつた。

(2)あなたの考える面白くなさとは

(1)とは逆に、面白くないことを自由記述で質問した。回答をまとめてみると次の通りである。

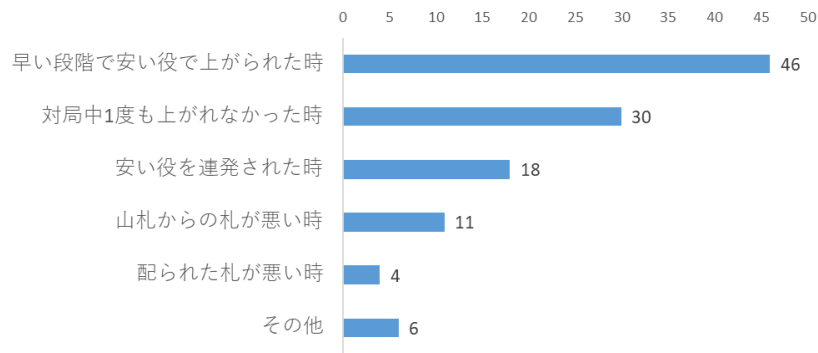


図 7-9 花札の面白くないこと(N=61)

- 早い段階で安い役で上がった時
- 対局中1度も上がれなかった時
- 安い役を連発された時

といったものがあった。その他にあたるものは「花見・月見が序盤に出た時」や「親権で流された時」などがあった。

(3)その他花札に対する考え

回答者に花札に対する考えを自由記述で尋ねたところ、41人が回答した。主なものを列挙する。

- 花見・月見の文数や採用の有無を考えるべきでは？(2枚で5文は安易すぎる)
- 主流はこいこいだが、ミツ(3人打ち)である花合わせも市販ゲームやflashゲーム、オンラインゲームになるといい。
- インフレルールを取り入れたら今の時代に合うのではないか。
- 役を減らしたり増やしたりしてバランスを取るのはいかがでしょうか。
- より普及できないか。

など、ルールに関するものやより一般に普及すると良くなるといった意見が目立った。

7.3.4 結果の分析

(1)年齢層ごとの面白さ

集計をした結果、10代と20代、30代から50代で考える面白さに傾向があることが判明した。以降便宜的に10代と20代を若年層、30代から50代を高年層と呼ぶこととする。図9は年齢層別の「あなたの考える面白さ」を示したものである。

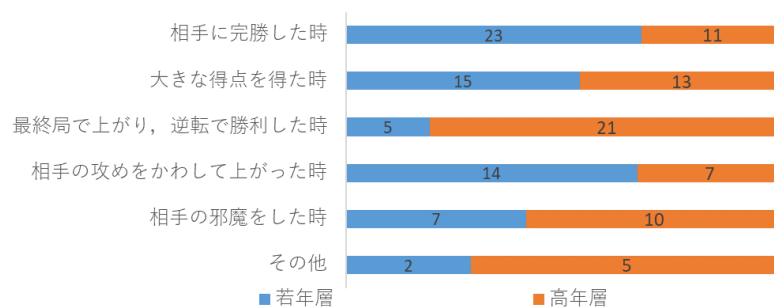


図 7-10 年齢層ごとの考える面白さ

若年層においては「相手に完勝した時」が最上位となり、強いことが面白いと考えていることが得られた。一方高年層では「最終局で上がり、逆転で勝利した時」が最上位であり、勝負の緊張感や勝負の過程に面白さの価値を見出すことを重視しているのが得られた。両者の最も重視する面白さの概念が違っているため、対象とする年齢層により演出する面白さが異なってくると言える。これをいかに表現するかが課題である。

また両年齢層ともに 2 番目であったものが「大きな得点を得た時」の面白さで、多くの役を成立させることを面白いと感じているのではないかと考える。従って役につながる札を場に捨てることで面白さを演出する近道であるが、切り札が不自然でありあからさますぎると面白くないと思われる。これを避け、自然な形で相手の人間プレイヤーに役を作らせる手法が必要である。

(2)年齢層ごとの面白くないこと

図 10 は年齢層別の「あなたの考える面白くないところ」を示したものである。

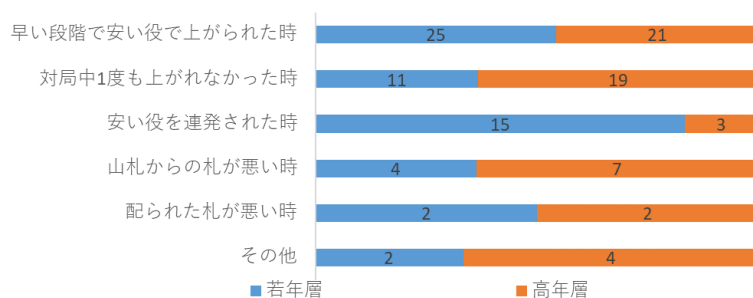


図 7-11 年齢層ごとの考える面白くないこと

両年齢層共に、「早い段階で安い役で上がった時」が最上位であった。従って面白くないと感じられないようにするため、早上がりを起こすような打ち方を避ける必要がある。この打ち方は以前構築したコンピュータプレイヤーでも見られることがあり、これのために評価が低かったのではないかと推測する。例えば、コンピュータプレイヤーが得点を±0に調整しようとする時、相手の人間プレイヤーが+3文の得点の場合に、コンピュータプレイヤーは3文程度の安い役でさっさと上がってしまうことがしばしば見られる。このため人間プレイヤーは面白くないと感じてしまうのではないかと考える。

面白くないことの 2 位は両年齢層で異なっており、若年層は「安い役を連発された時」、高年層では「対局中一度も上がれなかった時」であった。年齢層により避けるべき打ち方が違って来る。若年層に対しては「序盤で 1 文や 2 文で上がらず、こいをして得点を伸ばすようにする」打ち方が、高年層に対しては「適度に相手の人間プレイヤーの利になるような札を切る」ことで、その面白くなさを回避することはできるが、不自然な形で行ってはいけない。やはり自然な形での演出手法が必要である。

(3) 花札に対する考え方の違い

花札に対するコメントに関してまとめる。最も多かったのが「花見・月見」に関するもので、27 人（若年層 6 人、高年層 21 人）から意見があった。そのほとんどが「2 枚で 5 文は高すぎる、両方成立すると 3 枚で 10 文になりバランスが悪い」「序盤にこれが成立すると面白みがない」という意見であった。花見・月見は市販ゲームのほぼ全てに採用されており、製造元である任天堂の役一覧にも記載されている[1]正式な役であるが、バランスがとれていないというものが目立った。またこの意見を述べたのがほぼ高年層（77.8%）だったのも特徴的である。

次に目立ったのがルールや新しい役に関するもので、18 人から意見があった。例として「インフレルールを取り入れるのはどうか」「新役を取り入れて新しいものにするのはどうか」といったものである。ここで言うインフレルールとは「こい倍・7 文倍」というルールであり、「こいを 1 回するたびに得点が 2 倍になる」「得点が 7 文を超えると 2 倍になる」というルールである。このルールはどれほど得点が負けていようとどこからでも逆転が可能になるもので、ゲームを刺激的

で面白くする可能性があるものの、戦略が失われる面もあり一長一短である。また新役に関しても、広く普及しているものから変更・追加をするのは難しいと考える。ただ、同様に古くからあるゲームである麻雀においてはルールがかなり変わってきており、赤ドラを始めとするインフレルールが受け入れられるようになっている。従って花札においても今後の流れ次第ではルールが変わってくる余地もある。これらの意見は若年層から多く寄せられており、ここでも年齢層による考え方の違いが見て取れる。

(4)年齢層の傾向の分析

これまでの結果から年齢層ごとの傾向の分析を試みる。若年層においては「瞬間的な面白さを好む」という傾向があると考えられる。これは「相手に完勝した時」が最も高いことや、ルールへの意見で「インフレルールを採用したらどうか」が挙げられたことからこのように考えた。過程はどうあれ、大きな得点を挙げて勝てば面白いと思う傾向にあるのではないだろうか。また面白くないことの2位が「対局中一度も上がれなかった時」であり、適度に上がることで充足する傾向があると言えるだろう。一方高年層においては「過程を重視した面白さを好む」傾向があると考えられる。高年層は「最終局で上がり、逆転で勝利した時」を面白さの1番に挙げており、たとえ負けている状況であっても戦略的に得点を得て追いつき、最後には自分が勝利するという勝負の過程を楽しんでいるのではないかと考える。ルールに対する意見で「花見・月見」に関するものが多かったのも、過程を重視するため対局が序盤で安易に終わってしまうものを好まないからではないかと推測する。

(5) 今後に向けて

分析を行った結果、両年齢層で重視する面白さが異なっていることが判明した。従って演出すべき面白さが複数あることとなり、面白さの定義が複雑となる可能性がある。よって次のステップは本研究で得られたデータから両年齢層が満足するような定義付けを行うことが必要である。あるいは、全く別の視点を取り入れることも考えられる。それは「定義付けをせず、相手の打ち方から何を面白いと思っているかを推測する」という手法である。面白さの定義は言わば面白さの固定化であり、個々人の考える面白さが多岐にわたることが本研究で判明した以上、万人が満足する定義が出来ない可能性がある。そうなった場合に「相手の考えに合わせる」という手法を取る必要があると考える。

ある1つの先行研究では囲碁を対象とし、自然な囲碁を表現するために様々な研究を行っている[2]。一例として、人間らしさの表現に機械学習とモンテカルロ木探索を利用して「不自然すぎない悪手」を表現することに成功している。今後はより発展させて「指導」ができるAIを作成したいとしている。文献[2]の手法を参考とするならば「人間の価値観や好みの推定と迎合」が参考になると考える。つまり、相手の考える面白さを推定し、それに迎合するような打ち方をするものである。この手法なら本論文で示した「年齢層による面白さの違い」に対応できると考える。

7.4 生体情報を用いた面白さの要素の絞り込み

7.3 での面白さはある程度の方向性を示したものの、アンケート結果は人間プレイヤーの主観によるもので、面白さの定義についての議論をするためにはより客観的な指標が必要である。そこで、視覚で判断できる生体情報を用いることはできないかを検討した。現在では、生体情報は特定の情報であれば比較的簡単に計測が可能であり、数値を取得しグラフ化することにより目で判断できるようになる。生体情報の変化により何らかの発見があるならば、新たな知見が得られるのではないだろうか。

7.4.1 生体情報とは

生体情報とは、人間が発する身体の状態を表す情報のことである。通常はこれを測定して病気の有無や健康状態を見るものである。生体情報には数多くの種類があり、一例を図 7-12 に挙げる。図 7-12 は生体情報モニタにより計測できる情報で、よく知られているのは心拍数と血圧ではないだろうか。他例示した以外にも脳波や体温などが生体情報として計測されている。

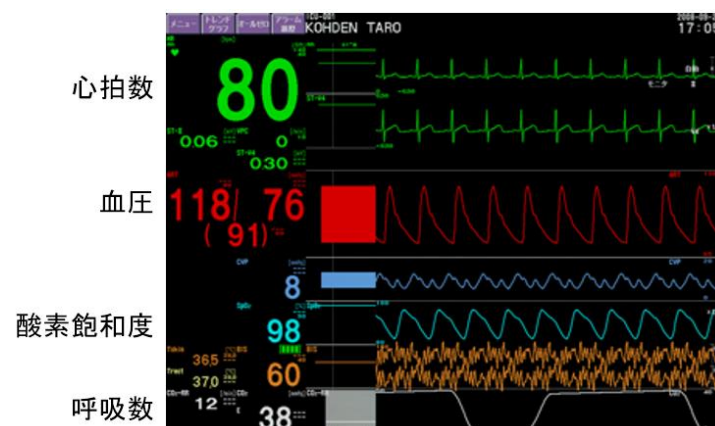


図 7-12 生体情報の例

心拍数や血圧は自身ではコントロールしにくいものであることから、花札をプレイする人間プレイヤーの生体情報を取得し心理的な変化から面白さを推定しようと試みた。重大な局面では緊張をしたり、思考を巡らせたりし、何かしらの心理的变化が生じるものである。従ってその変化を測定することにより「この局面はそのプレイヤーにとって緊張する局面である」ということが判明でき、即ちその局面が示すシチュエーションや要素が面白さであると言えるのではないかという仮説を立てた。本研究はこの仮説に基づきこの実験を行う。計測する生体情報は心拍数で、被験者へスマートウォッチを装着し、対局中の心拍数の変化を測定することにより面白さを推定するものである。

7.4.2 使用機材

心拍数を計るために、fitbit 社の charge2 というスマートウォッチを用いることとした。これは 1 秒単位で心拍数を取得することができ、局面における心拍数の変化を計るには十分であると考えられる。計測したデータは fitbit 社のサーバに蓄積され、時刻とともに記録が行われるので検証も容易であることも選択理由である。

7.4.3 実験方法

実験は被験者へスマートウォッチを装着させ、心拍数を計測しながら花札をプレイしてもらうことにより行う。またプレイ画面を録画し、局面と心拍数の関係を検証できるようにする。実験用のソフトウェアには 7.2.3 で開発した AI を組み込んだものを使用する。図 7-13 は実際のソフトウェア画面である。このソフトウェアは GUI を備え、札をクリックすることにより進行が可能である。本研究で

はさらにタッチパネルを採用し、札をタッチで選択することとした。このようにすることで、クリックするためにマウスを動かすラグが発生せず、よりリアルな変化が取得できるのではないかと考えた。なおここでは基礎研究を実施したため、被験者を一人に限って実験を行った。これはどのようなデータが得られるかを見極める必要があるという理由からである。

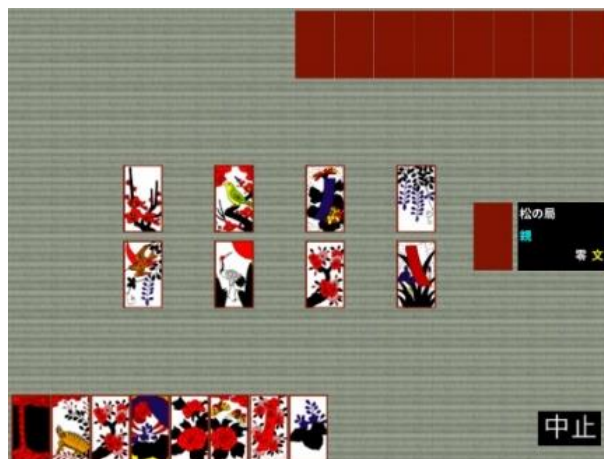


図 7-13 実験用ゲーム画面

7.4.4 実験結果

ここでは特徴のあった1局面を示す。図 7-14 はゲーム中の1局面、図 7-15 は図 7-14 の局面の始めから終わりまでの心拍数の変化である。心拍数のグラフは、縦軸が計測された心拍数、横軸が時間である。横に行くにつれて手番が進行していることを示す。丸印は図 7-14 の局面が現れた時点を意味する。図 7-14 の局面は、人間プレイヤーは花見で一杯の役を完成させようとしているが、コンピュータプレイヤーも三光を狙っている。人間プレイヤーの手札には桜の光札があり、場には桜の札はない。これより、人間プレイヤーが目指すことは桜の光札を取得することであり、コンピュータプレイヤーはそれを阻止して自分が光札を取ることである。

この局面において、人間プレイヤーの心拍数はゲームが進行するにつれて増加しており、手札が少なくなるたび緊張状の度合いを上昇させていると見られる。最終的に双方が役を完成させられず局が無勝負に終わった時がピークで、この局の間、人間プレイヤーは「自分はこの札を取れるのか？」を考え、緊張状態にあったと考える。であるから、緊張感をもたらす「札を取れるか否か」は面白さが感じられるものではないかと考えられる。また「札を取れるか否か」は、仮に観戦している人間がいるとすると見ていてわかりやすいシチュエーションである。従って「札の取り合い」という要素は花札の面白さの要素の一つ足り得る可能性がある。



図 7-14 ゲーム中の 1 局面



図 7-15 計測された心拍数

7.4.5 研究の現状と問題点

心拍数を計測する実験については、まず実験者が1人という状況である。そのためデータ数が不足しており、また多くの人数での実験を行っていないためデータの多様性もない問題がある。よって「この局面の要素が面白さだと言える」結論を出すのは尚早と言える。次に、測定している生体情報の種類の問題である。現在測定している情報は心拍数のみであり、他の情報は参照していない。より正確な結果を求めるには複数の生体情報を組み合わせることが重要ではあるものの、例えば脳波を測定するには脳波計が必要であり、機材の面で困難が伴う。今後、多くのデータを蓄積しようとするならば、実験で用いる機材をどのようにするか検討しなければならない。

7.5 まとめ

花札の AI を強くする研究を行った結果、強すぎる AI は面白くない AI であることが判明した。そこで、人間プレイヤーに面白さを与えるコンピュータプレイヤーは作成できないかを考え、角度の異なる 3 つの実験を実施した。

1 つ目は、独自に定めた面白さによって人間プレイヤーに面白さを与えられるかを検証する実験である。この実験には、強くする研究で用いた UCT を変形して適用し、定めた面白さに沿う札の選択をさせて人間プレイヤーを楽しませようとした。その結果、一部の面白さを満たすことができたものの、人間プレイヤーへ面白さを与えるところまでは行き着かなかった。

2 つ目は、AI の方向性を探るために花札の経験があるプレイヤーにアンケートを取り、そのプレイヤーが考える面白さについて尋ねたものである。結果を分析した結果、プレイヤーが考える面白さについての知見が得られた。しかし、完全に絞り込みができたとは言いがたかった。

3 つ目は、面白さの定量的な評価を目的に、プレイヤーの生体情報を用いて解析をしようとした。生体情報には様々な種類があるが、ここではプレイヤーの心拍数を用いて実験をした。人間の生体情報はコントロールしにくいものであり、ゲーム中の心拍数の変化を記録して解析することにより何らかの結果を得られるのではないかと考え、基礎的な実験を実施した。データ分析によって、一部の局面では心拍数に反応があることを確認した。

いずれの実験も、データ数の不足の問題を内包している。これは実験者を集めにくいことや、機材の関係上大規模な実験をしづらいことから来ている。今後は

この問題を解決し、より正確な定義付けを目指すような手法を構築する必要がある。

参考文献

[1] 任天堂 “花札の歴史・遊び方”

https://www.nintendo.co.jp/n09/hana-kabu_games/index.html

[2] 池田心: 人間プレイヤーを“楽しませる”囲碁プログラムの研究, 科学研究費助

成事業研究成果報告書, 基盤研究(C), 課題番号: 26330417, 2017

8 章 結論

ゲーム AI の研究が今盛んに行われており，研究が着手されているゲームも多い．本研究もゲーム AI を研究しようと考え，未だ着手されていない分野である花札の AI を研究したものである．以下，各章の総括を述べる．

1 章 序論

1 章では各章の概要を述べるとともに，研究背景を記述した．

2 章 ゲームの分類と意思決定手法

ゲーム AI を作成する場合はゲーム木を用いることが一般的である．しかし，探索する盤面が広大になるほどゲーム木の構築は難しくなる．2 章ではゲームの分類方法を挙げ，本論での分類方法やゲーム別の盤面の概算を示し，多くのゲームの盤面数が膨大であることを述べた．

次にゲーム木の概要を述べ，例を示しながら木の構築方法を提示した．それに付随してこの木を利用したアルゴリズムである Min-Max 法や α - β 法を解説した．また，新たなアルゴリズムとして機械学習を利用する例も見られ，機械学習の導入により大幅に実力が向上していることを示した．

3 章 ゲーム AI を強化する研究

3 章では，現在までになされているゲームの研究の内，本流であるゲーム AI を強化する研究を列挙した．チェッカー・チェス・囲碁の研究では，いずれのゲー

ム AI も強いこと，特にチェッカーは完全に解析されており，勝ち目がないことを述べた．また囲碁も棋力が上昇し，人類が勝てない領域に入りつつあることも記した．

4 章 本研究が対象とするゲーム

4 章では，本研究が対象とするゲームである花札の概要と共に，複数のバリエーションがある中，本研究で採用したこいこいの遊戯方法を解説した．

5 章 モンテカルロ木探索での強さの表現

モンテカルロ法をゲームに適用する場合は，シミュレーションを繰り返して良い手を求めるものとなる．また，良い手を集中的にシミュレーションすることでより正しい手を求めることができる．5 章ではゲームにおけるモンテカルロ法とその改良である UCT，多腕バンディット問題について解説し，木探索を行うための詳細な手順を記した．その後，花札の強いプレイヤーを作成することを目的とし，意思決定アルゴリズムに UCT を採用したプレイヤーを提案した．作成したプレイヤーの強さを評価するため，人間プレイヤーとの対戦結果を示した．この中で，UCT を採用したプレイヤーは人間プレイヤーに勝ち越しており，UCT が有効であるとの知見を得た．

6 章 ゲーム AI と面白さ

ゲーム AI の新たな研究の流れとして，人間プレイヤーに面白さを与えようとする研究が行われるようになってきた．6 章では面白さに関する研究を例示し，それ

らの研究がどのようになされたか、面白さを与えるための手法はどんなものであるかをまとめた。

7 章 面白さを表現するためのアプローチ

ゲーム AI の研究の第 2 段階として、6 章と同様に人間プレイヤーへ面白さを与える花札のコンピュータプレイヤーを作成できないかを検討した。この中で、手始めに 5 章で用いた手法である UCT を変形することで実現をしようと試みた。変形方法は、人間プレイヤーの得点を参照し、最終的な得点を ± 0 付近に調整して大勝ちと大負けを避けようというものであった。人間プレイヤーとの対局実験の結果、得点の調整に成功したものの、面白さの演出ができたとは言いにくかった。

続いて、正確な面白さの定義付けを目指すために、2 つの実験を行った。1 つは花札のプレイヤーへアンケートを実施し、プレイヤーが考える面白さについて尋ねることで、花札の何が面白いのかを絞り込もうとした。アンケート結果を分析したところプレイヤーの考える面白さがある程度見えてきたが、結果はプレイヤーの主観に頼っており、客観的な指標が求められることとなった。もう 1 つは人間プレイヤーの心拍数を計測し、反応のある局面にある要素を分析することで面白さのを見つけようとした。コンピュータプレイヤーとの対局実験により、対局中に心拍数の反応があることが判明した。しかし、この実験はデータ数が不足している問題が生じた。

8 章 結論

8 章では本論の総括を記述した。

今後の課題

強くする研究においては、より強いプレイヤーとすることが必要である。現状でもそれなりの強さを持っているが、まだ人間プレイヤーが勝つ余地がある。勝率を向上させることができれば、歯ごたえを求めるプレイヤーにとってはよいものとなるだろう。また切り札の改善もしなくてはならない。完成させられる役が絞られてくると、キーカードを捨てることが見られる。これを改善し、終盤まで札を保持するよう改良を行うことが不可欠である。

一方の面白さを与える研究については課題が多くある。第一にデータ数が不足している点、第二により客観的な指標が必要である点、第三に面白さを与えるためのよりよい手法を検討する点などであり、研究の完成には解決しなければならない問題が数多く残されている。

謝辞

本博士論文は北海道科学大学大学院工学研究科工学専攻博士課程での3年間の研究を総括したものであります。

本博士論文を執筆するにあたり、指導教員である北海道科学大学工学部情報工学科川上敬教授におかれましては、ゼミ等を通じて研究の方向性を示し、よりよい研究となるよう多大なご助言を賜りました。また北海道科学大学工学部情報工学科大江亮介講師には、日々の研究についてのご指導を賜りました。博士論文を作成できましたのは両先生のおかげであり、深く感謝申し上げます。また、本研究を行うに際して実験に協力頂きました皆様には感謝の念にたえません。この場を借りて深く御礼申し上げます。

最後に、本論文の執筆においてサポートして下さった全ての皆様に改めて御礼申し上げます、結びと致します。

業績一覧

投稿論文(査読あり)

- [1] Yuki Takaoka, Takashi Kawakami, Ryosuke Ooe: A Fundamental Study of a Computer Player Giving Fun to the Opponent, Journal of Computer and Communications, Vol.6, No.1, pp.32-41 (2018).

- [2] Yuki Takaoka, Takashi Kawakami, Ryosuke Ooe: Emergence of Fun Emotion in Computer Games -An experimental study on fun elements of Hanafuda-, Advances in Science, Technology and Engineering Systems Journal, Vol.3, Issue 6, pp.314-318 (2018).

- [3] Yuki Takaoka, Takashi Kawakami, Ryosuke Ooe: A Study on Estimating the Elements of Card Game Fun Using Biological Information, Journal of Computer and Communications, Vol.7, No.12, pp.129-135(2019).

会議論文(査読あり)

- [4] Yuki Takaoka, Takashi Kawakami, Ryosuke Ooe: A study on strategy acquisition on imperfect information game by UCT search, Proceedings of the 2017 IEEE/SICE International Symposium on System Integration, pp. 881-886. IEEE, Taipei, Taiwan (Dec. 11-14, 2017)

- [5] Yuki Takaoka, Takashi Kawakami, Ryosuke Ooe: A Fundamental Study of a Computer Player Giving Fun to the Opponent -Targeting Hanafuda, a Card Game in Japan-, Proceedings of the 7th International Conference on Kansei Engineering and Emotion Research 2018, pp. 441-450. Springer, Kuching, Malaysia (Mar. 19-22, 2018)

会議論文(査読なし)

- [6] ロボティクス・メカトロニクス 講演会 2017 in Fukushima

「花札のプレイヤーモデルに関する研究 -シミュレーションの深さによる面白さの変化-」

-Development of a player model in Hanafuda -Change in fun by depth of simulation-

2017年 5月 10日(水)～13日(土) ビッグパレットふくしま

- [7] 精密工学会 2019年度春季大会

「生体情報を用いたゲームの面白さを推定するための基礎研究」

2019年 3月 13日(水)～15日(金) 東京電機大学