

MM-NEAT を用いた一人麻雀における多目的最適化

Multiobjective Optimization using MM-NEAT in Single-player Mahjong

伊原 滉也 *¹ 加藤 昇平 *¹ *²
Ihara Koya Kato Shohei

*¹ 名古屋工業大学 大学院工学研究科情報工学専攻

Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

*² 名古屋工業大学 情報科学フロンティア研究院

Frontier Research Institute for Information Science, Nagoya Institute of Technology

In mahjong, there are several objectives in a hand to win the game. These are (1) to win quickly, (2) to get higher points, (3) to avoid losing points, and (4) to prevent other players from winning. These objectives are in the relationship of trade-off. In this research, we make an evaluating function of mahjong as a multiobjective optimization. On the other hand, mahjong requires multimodal behavior, because first place player have to win quickly and to avoid losing points, for last place player to get higher points in the final hand. Modular multiobjective NEAT (MM-NEAT) is a framework to evolve modular neural network. Each module defines a separate policy, and evolution discovers these policies and when to use them. In this paper, we focus on two objectives, (1) to win quickly, and (2) to get higher points, in single-player mahjong, which is a part problem of mahjong, and verify the efficiency of MM-NEAT.

1. はじめに

麻雀は1局ごとに、最大1人があがり、あがりによって点数を奪いあい、最終局が終了した後の得点で順位を決定するゲームである。

1局における目的は主に、

- 早くあがること
- 高い点数をあがること
- 失点しないこと
- 他のプレイヤーに得点させないこと

の4つである。これらの目的は複雑に関係しており、どの目的をどれだけ重視すべきであるか、それぞれの最適な重みは、点数状況や残り局数などによって動的に変化する。

例えば、暫定1位のプレイヤーは、早くあがること、失点しないことが重要になるが、4位のプレイヤーは高い点数をあがるのが重要となる。ただしどちらの場合もその他の目的を疎かにしてよいわけではない。

本研究では、まず麻雀の1局を多目的最適化問題とし、各目的に対してパレート最適な打牌選択をするエージェントのパレート解集合を獲得する。

また、麻雀ではあがりを目指すか、降りるか、一色手を狙うか三色手を狙うか、メンツ手を狙うかトイツ手を狙うか等、目的によって、相反する打牌選択になることが多く、状況に合わせたマルチモーダルな選択が必要となる。

そのため、MM-NEAT [Schrum 14] と呼ばれるモジュール構造を持つニューラルネットワークを進化により多目的最適化するフレームワークが有効に働くことが期待される。本稿では麻雀の多人数性を排除した一人麻雀において、早くあがること、高い点数をあがることの2目的にしばり MM-NEAT を用いたマルチモーダルな評価関数学習の有効性を検証した。

2. 1人麻雀

1人麻雀は、麻雀の多人数性を排除した部分問題である。34種類各4枚ずつの合計136枚あり、その中から14枚の牌で決

連絡先: 加藤昇平, 名古屋工業大学, 名古屋市昭和区御器所町,
052-735-5625, shohey@katolab.nitech.ac.jp

められた和了形をつくり、役に応じた点数を得る。1局の流れを以下に示す。

1. 牌をシャッフルし、山をつくる。
2. 山から順に13枚の牌を取り、手牌とする。
3. 山から牌を一枚取り、手牌に加える。
4. 手牌が和了の条件を満たすなら、その翻数と符に応じた点数を獲得し終了する。
5. 手牌から牌を一枚捨てる。
6. 捨てた牌が18枚なら終了する。そうでなければ3.に戻る。

採用した役を表1に示す。麻雀において、あがりの点数を高くするためには、翻数を高くすることが重要である(表2,3)。しかし役は翻数の高い役ほど、その成立難易度は高く、高得点を狙う場合どうしても聴牌までの効率を犠牲にすることになり、早さと高さはトレードオフの関係にある。

2.1 用語

和了 (ホーラ, あがり)

役と和了形が完成し、得点すること。和了 (アガ) ともいう。

和了形

和了できる形には決まりがあり、以下の3つのどれかを満たしていなければならない。

基本形 (面子手)

4面子1雀頭とも呼ばれる。その呼び名のとおり、雀頭と呼ばれる対子1組と面子4組で構成される最も基本的な和了形である。

七対子 (対子手)

対子を7組で構成される形。同種の牌4枚は2組の対子としてカウントされない。

国士無双

13種類の+幺九牌を各一枚ずつと、更にもう1つの幺九牌で構成される形である。

面子 (メンツ)

手牌の構成要素、基本形での和了を目指す場合はこれを集めることが一番の課題である。

表 1: 採用した役

役名	飜数	役名	飜数	役名	飜数
門前清模和	1 飜	全帯	2 飜	混一色	3 飜
平和	1 飜	混老頭	2 飜	清一色	6 飜
断ヤオ	1 飜	対々和	2 飜	四暗刻	役満
一盃口	1 飜	三色同刻	2 飜	大三元	役満
白	1 飜	三暗刻	2 飜	字一色	役満
發	1 飜	小三元	2 飜	大四喜	役満
中	1 飜	七対子	2 飜	小四喜	役満
三色同順	2 飜	二盃口	3 飜	九蓮宝燈	役満
一気通貫	2 飜	純全帯	3 飜	国土無双	役満

順子 (シュンツ)

同じ種類の数牌の 1,2,3 や 5,6,7 等, 数字の連続する 3 枚の組み合わせ.

刻子 (コーツ)

同じ牌を 3 枚そろえた組み合わせ.

対子 (トイツ)

同じ牌を 2 枚そろえた組み合わせ. あと 1 枚で刻子になる. 面子手なら雀頭として 1 つ, 対子手なら 7 つ必要となる.

塔子 (ターツ)

あと 1 枚で順子になる 2 枚の組み合わせ.

聴牌

手牌があと 1 枚で和了できる状態を聴牌 (テンパイ) という.

向聴数 (シャンテンすう)

聴牌まで, 最低何枚の牌が必要か表す値.

2.2 関連研究

水上らによって一人麻雀プレイヤを拡張して四人麻雀プレイヤを作成することの有効性が示された [水上 13].

海津らは一人麻雀において, パラメータ α の値を変更することで, 高ささと早さとバランスを調整する手法を提案した [海津 15].

2.3 多目的最適化

本研究における 1 人麻雀における目的は, (1) 早くあがること, (2) 高い点数をあがることの 2 つである. 本稿ではこれらを目的関数として定め, 1 人麻雀を多目的最適化問題として解くことで, パレート最適な麻雀プレイヤ集合を獲得する.

一人麻雀において, 不確定な要素は山の並び順のみである. 並び順を予想することは不可能なため, ある並び順に特化するのではなく, 全並び順 N 通りに対して同様の打ち方をしたときの, 平均的なあがりの早さと打点の高さを目的関数とする. 目的関数の式を以下に示す.

$$f_1 = \frac{1}{N} \sum_{k=1}^N (T - t(k)) \quad (1)$$

$$f_2 = \frac{1}{N} \sum_{k=1}^N P(k) \quad (2)$$

ここで, T は 1 局における最大ターン数, $t(n)$ は n ゲーム目終了時のターン数, $P(n)$ は n ゲーム目の獲得点数である.

3. MM-NEAT

MM-NEAT [Schrum 14] は, modular network と呼ばれるモジュール構造を持つニューラルネットワークを, module mutation

表 2: 点数表 (1 飜から 4 飜)

符\飜数	1 飜	2 飜	3 飜	4 飜
20 符	-	2000	3900	7700
30 符	1500	2400	4800	9600
40 符	2000	3900	5800	11600
50 符	2400	4800	7700	12000
60 符	2900	5800	9600	12000
70 符	3400	6800	11600	12000
80 符	3900	7700	12000	12000
90 符	4400	8700	12000	12000
100 符	4800	9600	12000	12000
110 符	-	10600	12000	12000

表 3: 点数表 (5 飜以上)

飜数	5 飜 (満貫)	6,7 飜 (跳満)	8,9,10 飜 (倍満)	11,12 飜 (3 倍満)	13 以上 (役満)
点数	12000	18000	24000	36000	48000

と呼ばれるモジュールを追加する突然変異と NEAT [Stanley 02] と NSGA-II [Deb 02] を用いて進化的に多目的最適化する手法である. NEAT とは, ニューラルネットの構造と重みを進化的に最適化する TWEANN (Topology and Weight Evolving Artificial Neural Network) の代表的な手法である.

3.1 Modular Network

モジュラーネットワークはモジュール構造を持つニューラルネットワークである. 図 1 のように, 出力層のニューロンはモジュールに分割される. それぞれのモジュールは preference neurons と policy neuron とよばれる出力ニューロンを持ち, 入力に対して preference neuron のニューロンの出力が最も大きいモジュールの policy neurons の出力が, ネットワーク全体の出力として使用される.

3.2 Module Mutation

Module Mutation とは, モジュラーネットワークに新しくモジュールを追加する突然変異である. 進化につれてモジュール数は増加していくことになるため, 初期解集団を単一のモジュールを持ったネットワークからはじめることにより, 適切なモジュール数を探索することができる. 新しいモジュールの Preference Neuron と Policy Neurons の初期化方法について, いくつかの手法が提案されているが, 本稿では最も代表的な MM(D) [Schrum 14] を用いる.

3.2.1 MM(D) [Schrum 14]

MM(D) では新しいモジュール内の preference neuron はランダムに選択された単一のモジュールからランダムな重みの入力

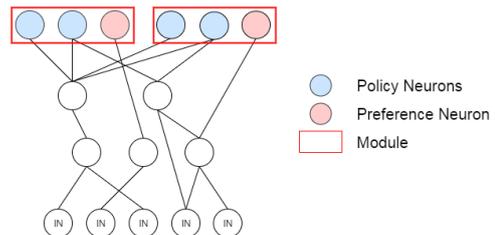


図 1: Modular Network (モジュール数 2 の場合)

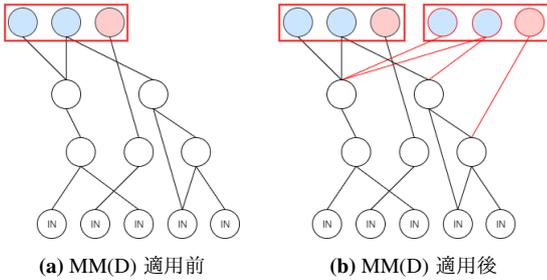


図 2: MM(D)

で初期化されるため、複製元のモジュールとは異なる状況で使用されやすくなる。また、policy neurons の入力ランダムに選択された既存モジュールの各 policy neurons を複製するように初期化される。これにより、モジュールが増加したことによる適応度減少を防げるため、後の世代にモジュールが増加した個体が生き残り、有用なモジュールを進化的に獲得できる。

3.3 NSGA-II [Deb 02]

多目的進化的アルゴリズムでは、多目的最適化を実現するために、パレート面へ探索を進めるための特別な個体の選択方法と集団の多様性を維持する方法が採用される。NSGA-II は、解のランキングに Goldberg によるパレートランク法を用い、ランク付けのためには高速非優越ソート (fast non-dominated sort) 法を用いる。個体選択には、パレートランクに基づいたタイプレックを持つバイナリ・トーナメント選択を採用し、適合度の割り当てには混雑距離を用いる。

3.3.1 アルゴリズム

NSGA-II のアルゴリズムの概要を以下に示す。

1. 初期集団の生成
大きさ N の初期個体集団 P をランダムに生成する。
2. ランク付けと混雑距離の割り当て
非優越ソートを用いて、 P の各個体にランクを付ける。
3. 子の生成
 P に対して、選択、交叉、突然変異の遺伝子操作を行い、大きさ N の子集団 Q を生成する。
4. 世代交代
非優越ソートを用い、親と子の全体 $P \cup Q$ の各個体にランクを付ける。また、各ランクごとに個体をソートし混雑距離を割り当てる。ランク上位 N 番目までの個体によって、集団 P を構成する。ただし、ランクが同じ個体については混雑距離が大きい個体を優先して選択する。
5. 3にもどる。

4. 一人麻雀の学習実験

麻雀の部分問題である一人麻雀において、MM-NEAT のマルチモーダルな学習が麻雀に対して有効に働くか検証する。例えば、通常手をあがるためのモジュールと、七対子をあがるためのモジュールが発見されることを期待する。

また、獲得したエージェント集合の性能評価のため、海津らの手法と比較する。一人麻雀の目的関数は式 (1),(2) の 2 つである。

4.1 実験設定

4.1.1 初期集団

初期個体はモジュール数 1 の modular network である。各 policy neurons は各入力ニューロンと結合し、preference neuron はランダムに選択された 1 つの入力ニューロンと結合する。重みは全てランダムに初期化される。

表 4: 入力情報

要素	次元数
手牌の牌ごとの枚数	5×34
通常形の向聴数	9
七対子の向聴数	7
順子の数	5
刻子の数	5
面子の数	5
対子の数	7
塔子の数	7
合計	215

4.1.2 ネットワーク

打牌後の手牌を入力、出力をその牌の評価値として、各牌ごとに評価値を計算し、一番評価値が高くなる牌を選択する。手牌の情報表現を表 4 に示す。入力情報は全て、ブール値で表現した。

4.1.3 適応度評価

個体の適応度評価は、それぞれの個体に実際に一人麻雀をプレイさせることで行う。麻雀では最善の手を選択し続けても聴牌できないこともあるほど、牌の並び順による点数のバラつきが大きいが、計算量の問題で、全並びの牌をプレイすることは現実的でないため、ランダムな並び順の山に対する 100 試行の平均値を適応度とした。また、聴牌してからは当たり牌が来るのを待つのみであり、和了できるかどうかや和了順目は、プレイヤーの戦略は関係なくそれ以降の牌の並び順にのみ影響される。そこで聴牌したところでプレイを終了とし、打点の高さは聴牌時のあがり期待点数で評価を行った。あがりの早さは聴牌の早さで評価した。

適応度関数の式を以下に示す。

$$f'_1 = \frac{1}{N'} \sum_{k=1}^{N'} (T - t'(n)) \quad (3)$$

$$f'_2 = \frac{1}{N'} \sum_{k=1}^{N'} P'(k) \quad (4)$$

ここで、1 個体あたりの試行数 $N' = 100$ 、最大ターン数 $T = 18$ とする。 $t'(n)$ は n 試行目に聴牌したターン数、 $P'(n)$ は n 試行目の聴牌期待得点を表す。

$$P'(n) = \sum_{i=1}^r \alpha_i \frac{a_i}{\sum_{i=1}^r a_i} \left(1 - \frac{N-nP_n}{N-nP_n}\right)$$

4.1.4 遺伝子操作

1 世代あたりの個体数は 100、最大世代数は 500 として実験を行った。遺伝子操作は、既存ニューロン間にランダムな結合追加 40%、既存結合間に新しいニューロンを追加 20%、MM(D)5% ネットワーク構造の交叉 (トーナメント選択)50% とした。

4.2 性能評価

各世代の適応度を図 3.4 に示す。 f'_1, f'_2 ともに適応度が上昇しており、MM-NEAT による多目的最適化が一定の効果を挙げていることが確認できる。

次に、500 世代目の優良個体に対して、ランダムな山を用いて 10000 局プレイさせた結果を表 6 に示す。評価は一人麻雀の強さの指標として和了率と一局あたりの平均点数、点数の高さの指標として和了時の平均点数、あがりの早さの指標として、和了平均順目によって行い、各評価項目に対して、最も優れた個体とその識別番号を抜粋した。

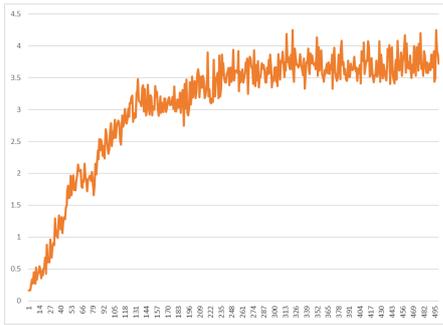


図 3: 各世代の f_1 の最大値

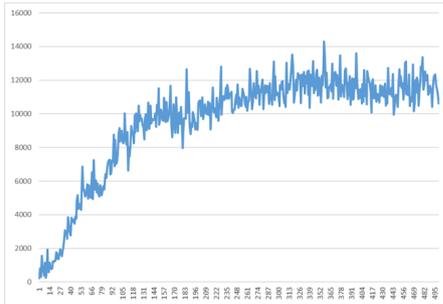


図 4: 各世代の f_2 の最大値

早さの指標として用いた和了平均順位において、海津らの手法よりもわずかだが優れた個体を獲得していることがわかる。

しかし、和了時平均点数においては、海津らは点数を上げる要素であるドラを用いているため単純に比較することはできないが、和了時平均点数において、大きく劣っている。これは、ニューラルネットの入力情報に役に関する知識をほとんど用いなかったためであると考えられる。役の知識を進化的に獲得することを期待したが、今回の実験で用いた入力情報では不十分であったと考える。

オンライン麻雀天鳳 [ten] における特上卓のプレイヤー（上位約 5%）の和了時の役複合割合を和了時平均点数最大の個体 (No.49950) の 10000 試行における役複合割合と比較した。表 7 は、一人麻雀に採用した通常形の役のうち上位 10 個とドラの複合割合を抜粋した。

表のとおりほとんどの役を十分にあげられていない。和了時の複合率上位の役をあげることは、平均的に点数をあげることにきわめて重要である。したがってこれらの役の知識を入力情報に入れることで全体平均点数、和了時平均点数の大幅な改善が期待できる。

また、和了率と全体平均点数においても海津らの手法と比べて、低い値となっているが、これは適応度評価に問題があったと考える。図 3,4 にみるとおり、個体の評価値の山によるバラつきが大きく、評価が十分正確であるとはいえない。これは、1 個体あたりの試行数を増加するか、配牌の時点の聴向数を制限するなどの工夫で改善できる。

5. おわりに

麻雀の部分問題における一人麻雀をあがりの早さと得点の高さに着目して多目的最適化問題とし、MM-NEAT を用いて評価関数の獲得実験を行った。

早さの指標において優秀な個体は獲得できたものの、得点の高さにおいて関連研究に劣る結果となった。

また実験に用いた手法には、役に対する知識不足や、ランダムな山による適応度評価のバラつきなどいくつか問題が見つ

表 5: 10000 局に対する優良個体の評価

MM-NEAT	和了率 (%)	全体平均点数	和了時平均点数	和了時平均順位
No.49959	13.13	402.0	3061.6	13.89
No.49968	12.57	447.6	3560.5	14.10
No.49950	11.38	415.1	3647.8	14.09
No.49948	10.64	356.9	3354.5	13.71

表 6: 10000 局に対する海津らの手法の最大値

海津らの手法	和了率 (%)	全体平均点数	和了時平均点数	和了時平均順位
$\alpha = 0.0$	17.51	693.0	3957.6	14.02
$\alpha = 0.1$	16.99	828.6	4877.0	13.95
$\alpha = 0.9$	13.16	931.5	7078.2	14.68
$\alpha = 1.0$	11.27	904.1	8022.4	14.90

表 7: 和了時の役複合割合の比較

役	天鳳特上卓 (%)	MM-NEAT (No.49950)
ドラ	38.31037	-
断幺九	23.09904	15.9
平和	19.44032	23.9
役牌 中	8.08352	2.02
役牌 白	8.06418	2.46
役牌 發	7.97840	2.19
混一色	5.05552	0.70
一盃口	3.86776	3.10
三色同順	3.61644	0.70
一気通貫	1.44501	0.08
混全帯幺九	0.90476	0.08

t

かった。今後は、これらの問題に着手する。また、今回の実験とは異なる、ドラを用いた環境での実験も予定している。

参考文献

- [Deb 02] Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE transactions on evolutionary computation*, Vol. 6, No. 2, pp. 182–197 (2002)
- [Schrum 14] Schrum, J. and Miikkulainen, R.: Evolving multimodal behavior with modular neural networks in Ms. Pac-Man, in *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, pp. 325–332ACM (2014)
- [Stanley 02] Stanley, K. O. and Miikkulainen, R.: Evolving neural networks through augmenting topologies, *Evolutionary computation*, Vol. 10, No. 2, pp. 99–127 (2002)
- [ten] オンライン麻雀天鳳, <http://tenhou.net/>
- [海津 15] 海津純平, 成澤和志, 篠原歩 他: 一人麻雀における打ち方を考慮した評価指標に関する研究, *ゲームプログラミングワークショップ 2015 論文集*, pp. 172–178 (2015)
- [水上 13] 水上直紀, 中張遼太郎, 浦見, 三輪誠, 鶴岡慶雅, 近山隆 他: 降りるべき局面の認識による 1 人麻雀プレイヤーの 4 人麻雀への適用, *ゲームプログラミングワークショップ 2013 論文集*, pp. 1–7 (2013)