

MM-NEATを用いた一人麻雀における 手牌評価関数の多目的最適化

伊原 滉也^{1,a)} 加藤 昇平^{1,2,b)}

概要：麻雀の1局における目的は主に、早く和了ること、高い点数を和了ること、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つであると考えられる。これらの目的は複雑に関係しあっており、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。そこで本研究では麻雀の一局を多目的最適化問題として捉え、パレート最適な手牌評価関数を持つ麻雀プレイヤーを獲得する。また、麻雀ではあがりを目指すか降りるか等、状況に応じたマルチモーダルな行動選択が必要となる。そこで、MM-NEATとよばれるモジュール構造を持ったニューラルネットの進化的最適化手法を、手牌の評価関数学習にとりいれる。本稿では、まず麻雀の部分問題である一人麻雀において、MM-NEATを用いて評価関数の学習実験を行い、MM-NEATの麻雀に対する有効性を検証した。

Multiobjective Optimization of Evaluation Function in One-Player Mahjong Using MM-NEAT

KOYA IHARA^{1,a)} SHOHEI KATO^{1,2,b)}

1. はじめに

麻雀は世界中の東・東南アジアコミュニティでプレイされる不完全情報ゲームである。麻雀は基本的には4人でプレイする。ゲームは局という単位に分割することができ、多くの場合1ゲームで4局か8局行われる。1局ごとに、4人のプレイヤーのうち1人のみが和了（あがり）ことができ、和了りによって点数を奪い合う。最終局が終了した後の得点で順位を決定する。

1局における目的は主に、以下の4つである。

- 早く和了ること
- 高い点数を和了ること
- 失点しないこと

- 他のプレイヤーに得点させないこと

これらの目的は複雑に関係しあっており、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、早く和了ること、失点しないことが重要になるが、最下位のプレイヤーは高い点数を和了ることが重要となる。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、面子（メンツ）手を狙うか対子（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ため、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、Modular Multiobjective NEAT (MM-NEAT [1]) 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的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

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

^{a)} ihara@katolab.nitech.ac.jp

^{b)} shohey@katolab.nitech.ac.jp

表 1: 本稿で用いた麻雀牌
Table 1 Tiles Used in This Paper.

		1	2	3	4	5	6	7	8	9
数牌	萬子									
	筒子									
	索子									
字牌	風牌	東	南	西	北					
	三元牌	白	發	中						

雀において、早く和了ること、高い点数を和了ることの2目的に着目し、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. 一人麻雀

本来、麻雀は4人でプレイするゲームであるが、一人麻雀は名前が示すとおり一人でプレイする麻雀である。そのため、プレイヤーは他のプレイヤーから点を奪われることがない。四人麻雀で和了を目指す場合、他人の持つ牌を考慮しなければ、一人麻雀とほぼ同じ打ち方になると考えられる。

本稿では表1に示した34種類の牌をそれぞれ4枚ずつの合計136枚使用する。プレイヤーは手牌を1枚ずつ交換していき、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2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3, 表4)。しかし役は翻数の高い役ほど、その成立難易度は高く、高得点を目指す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

本稿で使用する麻雀に関する用語を以下で説明する。

和了 (ホーラ, アガリ)

役と和了形が完成し、得点すること。和了 (アガ)る。

手牌 (テハイ)

プレイヤーが持つ13枚の牌。山から牌を引くと1枚捨てるまで一時的に14枚になる。このときの14枚が和了形を満たし、かつ役が1つ以上あれば和了できる。

面子 (メンツ)

手牌の重要な構成要素。一人麻雀では以下の2種類を合わせて面子と呼ぶ。

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。

塔子 (ターツ)

あと1枚で順子になる2枚の組み合わせ。例えば萬子の1,3は、あと萬子の2があれば順子になる塔子である。

和了形 (ホーラケイ)

和了するために必要な手牌の形。以下の3つのうちのどれかの形になっていないと和了れない。

基本形 (面子手)

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

萬子、筒子、索子の1,9と字牌すべての合計13種類を各1枚ずつと、13種類のうちのどれかをもう1枚の合計14枚で構成される形。

表 2: 一人麻雀に採用する役
Table 2 Yaku Used in One-player Mahjong

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

表 3: 点数表 (1 翻から 4 翻)
Table 3 Scoring Table (1-4 Han)

符\翻数	1 翻	2 翻	3 翻	4 翻
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 4: 点数表 (5 翻以上)
Table 4 Scoring Table (5 Han or more)

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

聴牌 (テンパイ)

あと 1 枚で和了できる手牌の状態.

向聴 (シャンテン) 数

聴牌まで, 最低何枚の牌が必要か表す値.

2.2 関連研究

松井らは打牌前の局面を入力, 打牌ごとの評価値を出力として, ニューラルネットを用いた打牌選択方法を提案した [2]. 教師データとして牌譜を用いていたが, 入力データに面子や雀頭を用いなかったことに加え, 教師データの偏りにより, 学習がうまく働かない問題があった. 水上ら [3] によって, 一人麻雀プレイヤーを拡張して四人麻雀プレイヤーを作成することの有効性が示されている. 海津ら [4] 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

さと早さとバランスを調整する手法を提案した.

2.3 多目的最適化

本研究における一人麻雀における目的は, 和了の早さと和了的点数の高さの 2 つである. 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であるため, 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.

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

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

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, $s(n)$ は n ゲーム目の獲得点数である.

多目的最適化では複数の目的を扱うため, ただ 1 つの最適解を求めるのではなく, パレート最適解の集合を求める. パレート最適解はパレート支配という解の優劣関係によって定義される. ベクトル $\vec{v} = (v_1, \dots, v_n)$ と $\vec{u} = (u_1, \dots, u_n)$ が以下の 2 式を満たすとき, \vec{v} が \vec{u} をパレート支配すると言う.

$$\forall i \in \{1, \dots, n\} : v_i \geq u_i, \text{ and} \quad (3)$$

$$\exists i \in \{1, \dots, n\} : v_i > u_i \quad (4)$$

この定義は, 解がある解をパレート支配するとき, 少なくとも 1 つの目的においてその解より優れており, 残りの目的すべてにおいてその解より劣っていないことを示す. ある解 \vec{x} をパレート支配する \vec{y} が存在しないとき, \vec{x} をパレート最適解であるという.

3. MM-NEAT

Modular Multiobjective NEAT (MM-NEAT [1]) 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 と呼ばれるモジュールを追加する突然変異と NEAT [5] を用いて進化的に最適化する手法である. 多目的最適化手法として Non-Dominated Sorting Genetic Algorithm II (NSGA-II [6]) が用いられる. 以下でこれらについて詳細を述べる.

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という単位に分割される. それぞれのモジュールは preference neuron と policy neurons と呼ばれる出力ニューロンを持つ. preference neuron はどの出力を使

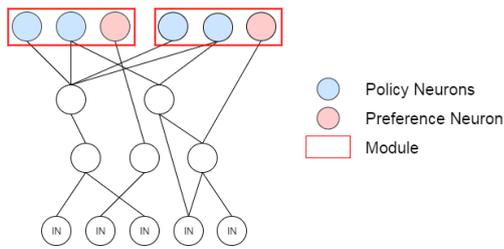
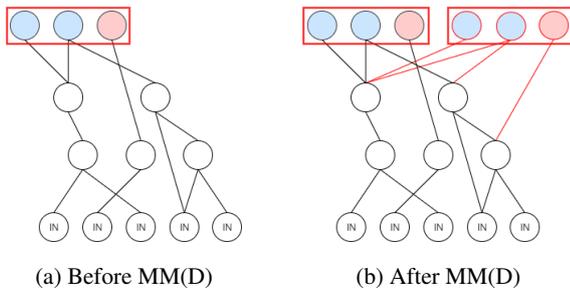


図 1: Modular Network (モジュール数 2 の場合)

Fig. 1 Modular Network (Two Modules)



(a) Before MM(D)

(b) After MM(D)

図 2: MM(D) の適用例

Fig. 2 An example of MM(D)

用するかを決定するニューロンである。policy neurons はそのモジュールの出力を決定するニューロンである。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 [7] 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、本稿では最も探索効率のよい MM(D) を用いる。

3.2.1 MM(D)[1]

Module Mutation Duplicate(MM(D) [1]) 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 の入力は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 NEAT [5]

neuroevolution はニューラルネットワークを進化的に最適

化する手法の総称である。neuroevolution の中でも、ニューロン間の重みだけでなくネットワークの構造も進化的に最適化するのが、Topology and Weight Artificial Neural Network (TWEANN) という枠組みである。Neuro-Evolution of Augmenting Topology(NEAT [5]) は、TWEANN で最も代表的な手法である。NEAT では初期個体は、隠れ層がなく入力層と出力層のみを持つネットワークである。ネットワークは weight mutation, link mutation, node mutation の 3 つの突然変異と topological crossover という交叉によって複雑化されていく。weight mutation は既存のリンクの重みを変化させ、link mutation は既存のノード間に新しくリンクを追加し、node mutation は既存のリンク間に新しいノードを追加する。これらの突然変異によって導入された全てのリンクとニューロンには、それぞれに固有の識別番号が与えられる。この識別番号は集団内の全個体で一貫しており、各ニューラルネットワークの遺伝子型はリンクとニューロンに加えこの番号を保存する。この表現により、異なる遺伝子間で共通の要素を識別番号で整理させることができ、交叉を効率的に行える。

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、通常手を和了するためのモジュールと、七対子を和了するためのモジュールが創発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 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。

4.2 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 5 に示す。入力情報は全て、ブール値で表現した。

4.3 適応度評価

また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。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。

表 5: 手牌の特徴量
Table 5 Features of A Hand

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

あがりの早さは聴牌の早さで評価した。また、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 N' 試行の平均値を適応度とした。また、適応度関数の式を以下に示す。

$$f'_1 = \frac{1}{N'} \sum_{k=1}^{N'} (T - t'(k)) \quad (5)$$

$$f'_2 = \frac{1}{N'} \sum_{k=1}^{N'} s^*(k) \quad (6)$$

ここで、 N' を試行数、 T を最大ターン数とする。 $t'(k)$ は k 試行目に聴牌したターン数、 $s^*(k)$ は k 試行目の聴牌期待得点を表す。一人麻雀では $T = 18$ である。

4.4 遺伝子操作

1 世代あたりの個体数は 300、最大世代数は 500 として実験を行った。遺伝子操作は、link mutation を 40%、node mutation を 20%、MM(D) を 5% トーナメント選択による topological crossover を 50% とした。

5. 実験結果

図 3、図 4 に各世代における f'_1, f'_2 をそれぞれ示す。 f'_1, f'_2 ともに世代を重ねるにつれて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上げていることが確認できる。

次に、図 6 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。表 7 に海津らの手法 [4] の結果を示し、これと比較する。早さの指標として用いた和了平均順目において、海津らの手法よりも優れた個体を獲得していることがわかる。

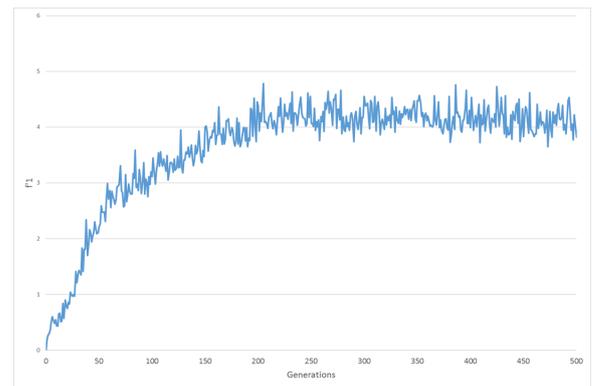


図 3: 各世代の f'_1 の最大値
Fig. 3 Largest Value of f'_1 in Each Generation

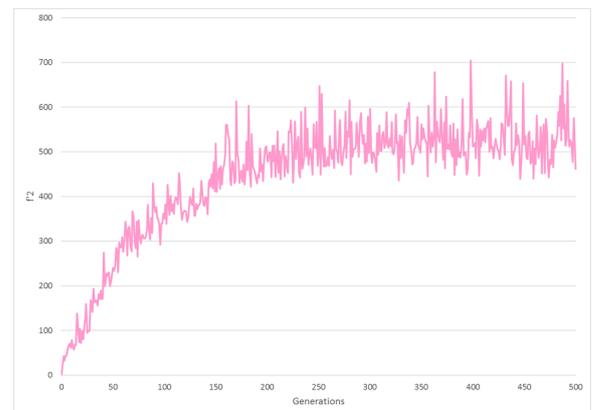


図 4: 各世代の f'_2 の最大値
Fig. 4 Largest Value of f'_2 in Each Generation

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と考える。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。

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をオンライン麻雀天鳳 [8] 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と比較した。表 8 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。ほとんどの役を十分に和了していない。これらの役を和了することは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組み込む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、図 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。

表 6: 10000 局に対する優良個体の評価

Table 6 Evaluation Values of Elite Individuals in 10000 Hands

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 7: 10000 局に対する海津らの手法の最大値 [4]

Table 7 Evaluation Values of Kaizu et al. in 10000 Hands

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 8: 和了時の役複合割合 (%) の比較

Table 8 Rates(%) of Yaku on Winning Hands

役	天鳳特上卓 [8]	MM-NEAT (No.49950)
ドラ	38.31	-
断幺九	23.09	15.90
平和	19.44	23.90
役牌 中	8.08	2.02
役牌 白	8.06	2.46
役牌 發	7.97	2.19
混一色	5.05	0.70
一盃口	3.86	3.10
三色同順	3.61	0.70
一気通貫	1.44	0.08
混全帯幺九	0.90	0.08

参考文献

- [1] Schrum, J. and Miikkulainen, R.: Evolving multimodal behavior with modular neural networks in Ms. Pac-Man, *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, ACM, pp. 325–332 (2014).
- [2] 松井一晃, 的場隆一ほか: ニューラルネットワークを用いた麻雀の打牌選択方法の提案, 研究報告ゲーム情報学 (GI), Vol. 2015, No. 8, pp. 1–5 (2015).
- [3] 水上直紀, 中張遼太郎, 浦晃, 三輪誠, 鶴岡慶雅, 近山隆ほか: 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).
- [4] 海津純平, 成澤和志, 篠原歩ほか: 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).
- [5] Stanley, K. O. and Miikkulainen, R.: Evolving neural networks through augmenting topologies, *Evolutionary computation*, Vol. 10, No. 2, pp. 99–127 (2002).
- [6] 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).
- [7] Schrum, J. and Miikkulainen, R.: Evolving multimodal networks for multitask games, *IEEE Transactions on Computational Intelligence and AI in Games*, Vol. 4, No. 2, pp. 94–111 (2012).
- [8] C-EGG: オンライン対戦麻雀 天鳳, <http://tenhou.net/> (2007).

6. 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改善を要する結果となった. 実験結果の考察から, 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の問題が見つかった. 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.