

2Dアクションゲームにおける 島モデルGAを用いた多様な振舞いの獲得

福嶋 良平^{1,†1,a)} 片寄 晴弘^{1,b)}

受付日 2016年12月21日, 採録日 2017年7月4日

概要: ビデオゲームの世界観を構築するうえで、Non-Player-Character (以下 NPC) にタイプの異なった性格付けを行うことはデザイン上の重要なタスクである。従来、これらの性格付け、それらが表出する NPC の振舞いについては、ゲーム制作者の経験的知識によって実装されてきたが、プレイヤーが見出す NPC の性格とは、与えられた制約条件下で最適化が行われた結果、典型化した振舞いに対して付与されたラベルとみることでもできる。本研究では、この考え方に基づき、島モデル遺伝的アルゴリズム (以下島モデル GA) を用い、先験的な情報を与えずに個性が表出する NPC の振舞いを獲得する手法を提案する。提案手法を 2D アクションゲームに対して実施したところ、慎重型、積極型などのように解される NPC の振舞いが獲得された。

キーワード: ゲーム AI, 島モデル GA, 2D アクションゲーム, 機械学習

Acquisition of Various Behaviors Using Island Model GA in 2D Action Game

RYOHEI FUKUSHIMA^{1,†1,a)} HARUHIRO KATAYOSE^{1,b)}

Received: December 21, 2016, Accepted: July 4, 2017

Abstract: In the video game, the characterization such as “bumpy type” or “discreet type” by Non-Player-Character (NPC) is extremely important in user experience design. Traditionally, these personality characterizations have been made by empirical knowledge of experienced game makers, but the personality in NPC is the behavior optimized by collective intelligence, that is, under a given constraint, it can also be thought as a label given to a certain strength and secured behavior. In this research, based on this hypothesis, we propose a method to acquire NPC behavior without a priori information using island model genetic algorithm (island model GA). Here, when the proposed method was applied to the 2D-action-game, several behavior patterns were acquired without giving a priori knowledge by the island model GA.

Keywords: game AI, island-model-GA, 2D-action game, machine learning

1. はじめに

ビデオゲームにおいて、Non-Player-Character と呼ばれる人間プレイヤーが操作をしないキャラクタ (以下 NPC) は

ゲーム AI と呼ばれる思考ルーチンにより制御されており、その自動構成は、エンタテインメントコンピューティング領域の主要研究テーマの 1 つである。強い AI の構成を目的とした研究は、ボードゲームの分野において先導的に実施され、「あから 2010」が将棋のプロ棋士に勝利するなどの成果があがっている [1]。この事例が示すように、強い AI の構成という学術的なゴールが達成されつつあるが、代わって、特に、ビデオゲームのようにユーザが楽しむことを主目的とした領域において、「人間らしさ」や「自然な弱さ」の構成に関する関心が大きく高まりつつある [2], [3], [4], [5]。

¹ 関西学院大学

Kwansei Gakuin University, Sanda, Hyogo 669-1337, Japan

^{†1} 現在、関西学院大学大学院理工学研究科

Presently with Graduate School of Science and Technology, Kwansei Gakuin University

a) erv95460@kwansei.ac.jp

b) katayose@kwansei.ac.jp

ビデオゲームにおいてタイトルの世界観を構築するうえで、登場する NPC タイプに性格付けを行うことは、デザイン上の主要タスクとして位置付けられる。従来、ビデオゲームの制作現場では、人手によって、NPC の性格付けが決定され、それらが適切に表出されるよう振舞いが設計されてきた。近年、ゲーム規模は拡大傾向にあり、ゲーム内に登場する NPC の種類やとりうる行動が多様化し、その設計には大きなコストがかかるようになってきている。この問題に対して、生物学的制約を取り入れた人間らしい振舞いを持つゲーム AI の獲得 [3] や、GA を用いた人間プレイヤーの強さに適応するゲーム AI の獲得 [4] など、特定の振舞いや強さを持つゲーム AI の自動獲得に関する研究が提案されてきた。これら提案手法によって、「自然な弱さ」を表出するゲーム AI の自動構成の筋道は示されたが、異なった複数の振舞いの獲得や、それらの振舞いがプレイヤーに与える印象の検討を行った研究は存在していない。

本論文では、性格特性などの先験的な情報を与えることなく NPC の多様な振舞いを獲得する手法として、鳥モデル遺伝的アルゴリズム（以下鳥モデル GA）を用いる手法を提案し、2D アクションゲームを対象としてその効果を検証する。

2. 関連研究

従来より、ゲーム AI の自動構成を目的として、様々なアプローチでの研究が進められてきたが、機械学習技術の進展、マシンパワーの増大により、強さの追求という研究目標は、現時点で、ほぼ達成されつつある [6], [7], [8], [9]。これに代わり、最近では、プレイヤーを楽しませることを目的とした研究への注目が高まりつつある [2], [3], [4], [5]。

人間プレイヤーの好敵手となるゲーム AI の構成を目指したものとしては、オセロゲームを対象とした上田らの研究があげられる [4]。この研究では、最適行動を選択するゲーム AI を用意しておき、行動選択の一致率から人間プレイヤーの強さを推定して、遺伝的アルゴリズムにおける適応度関数に組み込んで参照する。この操作により、一般のプレイヤーに対しておおよそ五分の勝率となるゲーム AI を構成できることが報告されている。

一方、人間らしい行動選択の自律的な獲得を目指したものとしては藤井らの研究があげられる [3]。この研究では、強化学習を用いたアクションゲームのゲーム AI の振舞いの獲得において、「疲れ」や「間違い」などの生物学的制約を導入した強化学習を用いることで、一定のうまさを確保しつつ、「無理をしない」「戸惑った様子が見える」など、人間らしいと解される振舞いが獲得できることを報告している。

ゲーム AI の行動選択に特定の個性を持たせることを目指した研究としては、First Person Shooting ゲームを対象とした Esparcia-Alcázar らの研究があげられる [10]。この研究では、あらかじめ NPC がとりうる行動を「攻撃的行動」

「保守的行動」の 2 つに分類し、それぞれの視点での最適化問題に GA を採用することで、NPC の振舞いに違いを生み出すことに成功している。

以上のように、プレイヤーのユーザエクスペリエンスデザインに留意したゲーム AI に関する研究は大きく進展しつつあるが、「人間らしさ」や「個性」の自律的な構成については、解決すべき事項が多く残されている。文献 [10] に代表されるように、複数の「個性」の獲得に成功した例もあるが、プログラマが前もって「攻撃的行動」「保守的行動」などの独立した評価基準を用意する必要がある。また、機械学習により獲得された NPC の振舞いについての印象の検討は十分になされているとはいえない。

3. アプローチ

本研究では、性格特性などの先験的な情報を与えることなく、複数の「個性」を有した NPC を自律的に構成することを目的とする。一定の強さ（上手さ）の確保、多様な振舞いの一括獲得という目標に際し、鳥モデル遺伝的アルゴリズム（以下鳥モデル GA）を採用し、獲得された NPC の振舞いについての印象評価を実施する。

ビデオゲームの世界観を構築するうえでは、NPC に「猪突猛進型」あるいは「慎重型」などのようにタイプの異なった性格付けを行うことはゲームデザイン上の重要なタスクである。ゲームの制作現場においては、NPC の性格付けの設定が行われ、人手によって、それらが表出するよう振舞いの実装がなされている。しかし、人間プレイヤーが感じ取る NPC の性格とは、先見的に存在した概念というより、典型化した振舞いに対して付与されたラベルと考えた方が自然である。この考えに従うと、NPC を構築するうえで、機械学習により典型的な行動パターンを獲得することができれば、性格特性を与えていなくても、人間プレイヤーから見て、個性や性格が知覚されると思われる。本研究ではこの考え方に従い、性格特性などの先験的な知識を与えることなく、一定の強さを確保するよう最適化を進めた NPC の振舞いのクラスターを複数導出し、それらがどのような「性格」として解されるのかの検証を進めていく。

学習のための入力データと出力データのペアを与えずに最適化を行う手法としては、TD 学習や Q 学習が知られているが、本研究では、進化アルゴリズムの一種である GA を採用する。GA においては、適応度関数に目標とする NPC の強さを組み込むことで、強さをある程度制御しつつ、最適解を得ることができる。また、振舞いを求めるべきパラメータセットとすることで、一定の強さを担保しつつ NPC の振舞いを獲得できる。しかし、通常の GA では、早期収束の影響などの特性により、得られるパラメータセットは単峰となる。そのため、複数の振舞いを獲得するには適さないといえる。そこで、多峰解を導出することに長けている鳥モデル GA を活用する。

対象とするゲームには、強さの違いや振舞いの多様性が表出しやすいことが求められる。本研究では、2D アクションゲームの基本要素を有する独自によって開発されたゲームを用いての実験を実施する。獲得された NPC の特性の分析に際し、まずは、行動選択ログをシミュレーションによって取得し、ガウス混合分布（以下 GMM）を用いたクラスタリングにより、数学的な分析を実施する。続いて、人間による主観評価実験を実施する。

4. 実験用 2D アクションゲーム

本章では、提案手法の実装対象として採用する 2D アクションゲームの概要と実験用プラットフォームのデザインについて述べる。

4.1 2D アクションゲーム

2D アクションゲームとは、最もポピュラーなリアルタイム性を有するゲームジャンルの 1 つであり、2D 空間上に表示されるプレイキャラクターの移動行動、攻撃行動、あるいは、ゲームタイトルごとに設定された特殊行動をプレイヤーが選択・操作することでゲームが進行していく。2D アクションゲームは、さらには、対戦型やステージ攻略型などに分けることができるが、対戦型の 2D アクションゲームの多くでは、プレイキャラクターを操作して、制限時間内に特定の敵キャラクターを倒すことが目的となる。勝ち負けが明確であり、プレイヤーの強さや上手さが顕著に現れる。また、ゲーム状態は無数に存在し、操作タイミングを含めて行動選択の傾向の違いが顕著に現れる。

4.2 実験用ゲームのプラットフォーム

本研究においてはゲーム内部の情報にアクセスできる必要がある。市販ゲームでは、この目的に対応することができないため、ここでは、独自に開発した 2D アクションゲームを用いて実験を実施する。開発したゲームのプレイ画面を図 1 に示す。

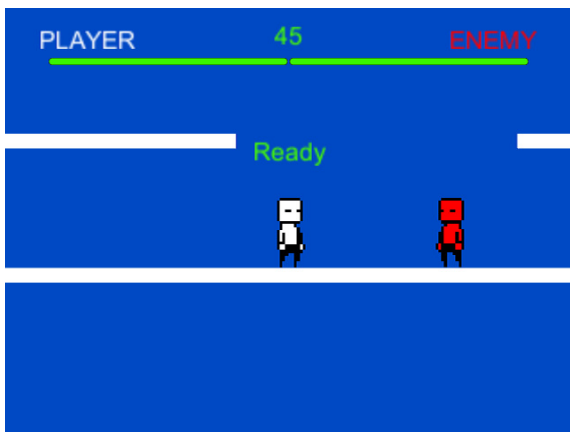


図 1 開発したゲームのプレイ画面
Fig. 1 Play screen of developed game.

このゲームは 1 体の敵キャラクターと戦うゲームであり、制限時間内に敵キャラクターを倒すことをゲームの目的とする。各キャラクターはゲーム中に、移動・近接攻撃・遠距離攻撃・防御のいずれかの行動を行うことができる。移動行動は、歩き・走りの 2 種類が存在する。各キャラクターのゲーム開始時の体力は 100 とし、近接攻撃・遠距離攻撃の攻撃が敵キャラクターに当たった場合にはそれぞれ体力を減少させる。減少量は近接攻撃は 10、遠距離攻撃は 2 とし、敵キャラクターの防御中に攻撃行動があつた場合のダメージ量は近接攻撃は 2、遠距離攻撃は 1 とする。制限時間内にどちらかのキャラクターの体力が 0 となるか、制限時間を超過するとゲームが終了する。ゲーム終了時点の残り体力が多いキャラクターが勝利となる。

4.3 ゲーム中のキャラクターの動作制御

今回使用する 2D アクションゲームでは、市販のビデオゲームでも活用されているゲーム AI に準拠した方法である、BehaviorTree を用いて NPC の制御を行う。BehaviorTree は木構造で NPC の行動を記述する手法の一種であり、ゲームの局面状態と、BehaviorTree に設定されるパラメータから NPC の行動を決定する。BehaviorTree を用いたゲーム AI を自動獲得するためには、行動決定の基準となるパラメータセットを最適化する必要がある。

NPC の多数の振舞いを獲得するためには、ゲーム AI の基本構造自体に多様な振舞いをとりうる余地が必要となる。今回構築する BehaviorTree では、行動選択の幅を確保するためのいくつかの工夫を施す。図 2 は実装した BehaviorTree の全体像を表している。図 2 中、a ではキャラクターがどれだけ連続して行動選択を行っているかを、条件式 (1) を用いて判定する。ここで $flame_n$ は、ゲーム側から観測される、キャラクターが連続で操作されているフレーム数を表し、 $flame_{max}$ は、BehaviorTree 側で設定される連続操作を許容する最大のフレーム数を表す。 $Score_1$ が 1 以上となる場合は、キャラクターは「停止行動」をとり、それ以外の場合は条件判定 b へと進む。

$$Score_1 = flame_n / (flame_{max} * 10) \quad (1)$$

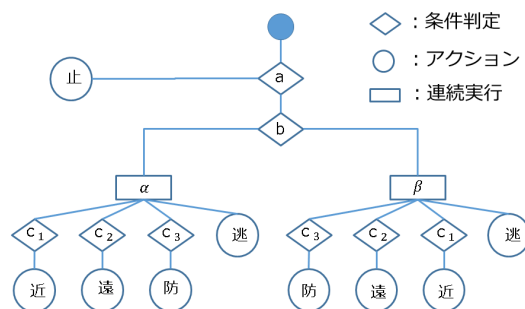


図 2 BehaviorTree の全体像
Fig. 2 Overview of BehaviorTree.

図 2 中, b では, 現在の状況が有利であるか否かを, その時点での残り体力の差によって導き出している (式 (2)). P_{hp} および E_{hp} はそれぞれ, BehaviorTree の探索時点でのプレイキャラクタ・敵キャラクタの残り体力であり, ゲーム側から観測される. $W_{E_{hp}}$ は, 敵キャラクタの体力に対する重みパラメータであり, BehaviorTree に対して設定される. $Score_2$ が 0 以上となれば, その状況を有利と判定し, α へ進み, それ以外の場合は不利と判断して β へ進む. α および β 以降では, それぞれで「近接攻撃」などの, どの行動を選択するか, 判定する優先順位を入れ替える. 有利な場合は, 「近接攻撃」「遠距離攻撃」「防御行動」の順に判定を行い, 不利な場合には「防御行動」「遠距離攻撃」「近接攻撃」の順に判定を行う. どちらの場合にも, いずれの行動選択も行われなかった場合は「逃避行動」を行う. このようにすることで, 有利な場合と不利な場合でそれぞれ違った状況判断を行うようにし, 多様な行動選択をとりうる余地を見出す.

$$Score_2 = P_{hp} - W_{E_{hp}} * E_{hp} \quad (2)$$

キャラクタの各行動を実施するか否かは, 条件式 (3) によって求める. Dis はキャラクタが移動できるステージの両端の距離を 1.0 としたときの, プレイキャラクタと敵キャラクタの距離を表す. Dir は, 敵キャラクタの向きを表し, プレイキャラクタに向かっている場合は 1, 背を向いている場合は -1 とする. $state$ は敵キャラクタの現在の行動に対する重みパラメータであり, 歩き行動, 走り行動, 停止および防御行動のいずれを行っているかで特定の数値を返す. 歩き・走り行動には, 単純な移動行動に加え, 歩きながら (走りながら) 行う近接攻撃・遠距離攻撃も含まれる. 歩き行動, 走り行動, 停止および防御行動はそれぞれ $state_1, state_2, state_3$ の変数で表し, BehaviorTree 側で 1 から 100 の値を設定する. $judge$ はプレイキャラクタが, 近接攻撃・遠距離攻撃・防御行動のいずれかを行うと判定する際のバイアス値として BehaviorTree 側で設定するパラメータである. この値は, 図 2 中, c_1, c_2, c_3 それぞれの条件判定に対し独立したパラメータ $judge_1, judge_2, judge_3$ を設定し, プレイキャラクタにおける最終的な行動選択の判断規準として取り扱う. 実際には, $Score_3$ の計算結果が, ± 0.1 以内となった場合にのみ, 対応する行動を実施し, それ以外となった場合には, 次の行動の実施を判定する.

$$Score_3 = \frac{abs(Dir * Dis + state_n/100)}{2.0} - \frac{judge_c}{100} \quad (3)$$

5. 島モデル GA の構成

本研究では, 島モデル GA を, BehaviorTree における行動を決定するための重みパラメータの最適化に活用する. ここでは, 島モデル GA の遺伝子情報として, BehaviorTree

のパラメータセットを与え, 適応度関数に目標とする NPC の強さを取り込み, 学習を実施する. 島モデル GA の実施においては, 適応度関数における NPC の強さを計測するために, 生成される個体についてゲームをシミュレートさせる. 本章では, 島モデル GA の概要と, 個体の生き残りを決定するためのシミュレーション方法について述べる.

5.1 島モデル GA の概要

島モデル GA とは, 多次元パラメータの最適解を探索的に導き出す手法である GA の一種である. 通常の GA は最適化対象となるパラメータセットを個体を与え, 個体の交叉・評価・世代交代を繰り返すことで, 探索的に最適解を導出する. 島モデル GA では最適化するパラメータ空間を, 複数の島に区切り, 島ごとに個体の交叉・評価・世代交代を繰り返すことで, 島ごとの最適解を導き出す手法である. しかし, 島内部だけでの世代交代では, より優秀な個体が見つけれない場合もあるため, この手法では一定世代ごとに, 移住と呼ばれる島間の個体の交換を行う. 移住を行うことで, 優れた個体を持たない島は, 優れた島に淘汰されるため, 個体の収束を促すことができる. 島モデル GA の全体的な手順については, 図 3 に記す. 島モデル GA では, 複数の極所解を導出できる特性があり, 「NPC の振舞い」を求めるべきパラメータセットととらえることで, 得られる極所解ごとに異なった振舞いを有していることが期待できる. また, 適応度関数に目標とする NPC の強さを取り込むことで, 強さをある程度制御することができる. ここでは 2D アクションゲームにおける NPC の獲得のために実装した島モデル GA の概要について記す.

表 1 に今回実装した島モデル GA の詳細を示す. 島モデル GA の遺伝子情報としては, 式 (1) から式 (3) までの BehaviorTree 内で用いられる評価関数における, $flame_{max}, W_{E_{hp}}, state_1, state_2, state_3, judge_1, judge_2, judge_3$ の 8 つのパラメータを与える. これらの重みパラメータの範囲は 1 から 100 とする. 世代交代モデルには佐藤らにより提唱されている MGG-best2 モデル [11] を用

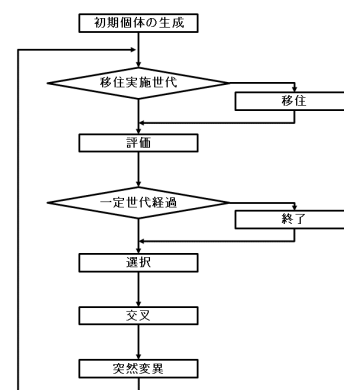


図 3 島モデル GA の手順のフロー図
Fig. 3 Flow diagram of island model GA.

表 1 島モデル GA の詳細
Table 1 Detail of island model GA.

項目	備考
世代交代モデル	MGG-best2
交叉法	BLX- α
遺伝子情報	BehaviorTree 内の 8 次元パラメータ
各遺伝子の範囲	1~100
各遺伝子の初期値	範囲内の正規乱数により決定
島の数	25
一島あたりの個体数	4 個体
世代交代の回数	40 世代
移住の間隔	3 世代ごと
移住対象	各島の最も優良な個体
移住先	各島の重心の最短巡回路に従う

いる。交叉法としては、遺伝子情報が実数値であるため、Eshelman らにより提唱されている、BLX- α 法 [12] を用いる。GA における突然変異は、BLX- α 法における α 値によって再現する。ここでは、 α 値を 0.5 とする。1 世代あたりの個体数は 100 個体とし、交叉の際には新たな個体を 10 個体生成する。

本研究では、NPC の多数の振舞いの一括獲得を目標としているが、獲得する NPC のパターン数については明らかになっていないため、極力多くの島を構成する必要がある。そこで、島 1 つあたりの個体数を 4 個体とし、1 世代あたりの島の数を 25 個とする。島モデル GA 特有の作業である「移住」については、島 1 つあたりの個体数が少ないため、島内での早期収束が懸念される。そのため、移住の間隔は 3 世代ごととする。移住の際、個体は各島におけるその最も優秀な個体を選択する。移住先は、個体の遺伝子情報重みパラメータによって形成される島の重心とする。この際、移動させる個体の移動距離の総和が最小になるように最適化処理を施す。なお、最大世代数は 40 世代とする。

個体の適応度については、プレイキャラクターの強さを取り込む。プレイキャラクターの強さは、試合の終了時点での各キャラクターの残り体力の差で表し、式 (4) で求める。 S の値が高ければ高いほど、敵キャラクターからの攻撃を受けず、ゲームを有意に進めることができたことを意味し、低いほど敵キャラクターに対して有効な攻撃を与えることなく、完敗したことを表す。適応度関数については、プレイキャラクターの強さ S を基に式 (5) により求める。この式において、 T_{value} は NPC の強さの目標値を表している。ゲーム 1 試行終了時点での各キャラクターの残り体力の差が目標値にどれだけ近いかを、式 (5) で計測しており、 f の値が高いほど良い個体となる。学習の目標値としての値を島モデル GA の実施ごとに調節することで、特定の強さを持った AI を導出する。

$$S = Last_{P_{hp}} - Last_{E_{hp}} \quad (4)$$

$$f = 100 - abs(T_{value} - S) \quad (5)$$

5.2 島モデル GA における個体の評価のためのシミュレーション方法

島モデル GA によって生成される個体については、その個体の良し悪しを判定し、次世代に生き残る個体を決定する必要がある。個体の良し悪しを判定する方法については、生成された個体についての強さを計測するため、ゲームをシミュレートし、その結果から適応度を求める方法を採用した。シミュレーションの手順は以下のとおりである。

- (1) 各キャラクターの位置・体力を初期状態とし、制限時間を 45 秒とする。
- (2) 試合を開始する。
- (3) 終了条件を満たすまで試合を継続する。
- (4) ゲームが終了した時点で適応度を求める。
- (5) シミュレーションを終了する。

シミュレーションは島モデル GA により生成される個体につき 1 回ずつ行う。敵キャラクターには、「近接攻撃」「遠距離攻撃」「防御」「逃避」「停止」の 5 つの行動を一様乱数に基づいてランダムに行う AI を使用する。ゲームの終了条件としては、制限時間が 0 となるか、プレイキャラクター・敵キャラクターのいずれかの体力が 0 となった時点とする。ゲームが終了した時点でのプレイキャラクター・敵キャラクターの残り体力から、式 (5) を計測し、その個体の適応度とする。

6. 評価実験

本章では、島モデル GA によって導出されたゲーム AI の「強さの制御」と「生成された振舞い」の両視点での評価について述べる。「生成された振舞い」については、NPC の行動選択ログの比較と主観評価実験の 2 つの視点から評価を実施する。前者については、NPC の行動選択状況に対して GMM を用いたクラスタリングを実施し、典型的な行動選択を、数学的なアプローチで検証する。後者では、前者の検証により得られた典型的な行動選択パターンが観察者にどのような印象を持たれるのかについて検証する。

6.1 NPC の強さの制御に関する評価

本節では、島モデル GA の動作確認、および、目的とする強さの可制御性を検証を実施する。この検証を行うため、最強・最弱を表す NPC の評価の目標値をそれぞれ、100、-100 と設定し、6 段階のレベルで、目標値どおりに適応が進んでいるかを検証する実験を実施する。

上記の目標値に対して島モデル GA を適用し、得られた個体の遺伝子情報のベクトル空間について、BIC 基準によって最適なクラスタ数を求めたところ、目標値を -60、20、60 として得られた個体は 5 クラスタに、-100、-20 として得られた個体は 7 クラスタに、100 として得られた個体は 8 クラスタに分類され、島が形成されることが確認された。

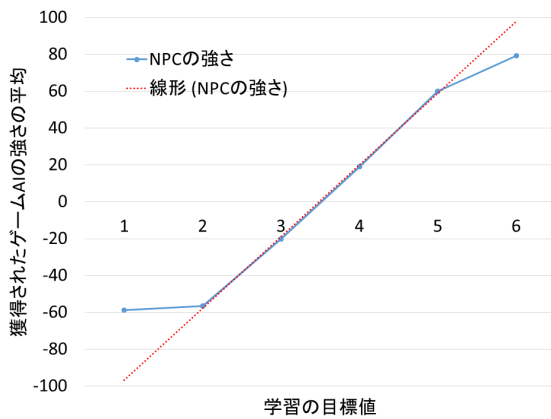


図 4 各学習における学習結果の NPC の強さの平均

Fig. 4 Average of NPC's strength at each learning result.

続いて、各目標値ごとに得られた NPC の強さを、式 (4) により求めた結果を図 4 に示す。図中の実線は、目標値ごとに得られた NPC の強さの平均を表し、破線は得られた NPC の強さのうち、最強・最弱となることを目的として学習した結果を除いて一次フィッティングを行った直線である。この結果より、最強・最弱以外の目標値に対しては、ほぼ目標どおりの強さになるように学習が進行していることが分かる。最強・最弱を目標とした学習結果については目標値に届かない結果となった。最強については、ゲームシステム上、プレイキャラクターが敵キャラクターに攻撃を与えた際に反撃を防ぐ手立てがなかった、また、最弱については、BehaviorTree の構成上の制約として、完敗する状況、つまり、敵にダメージを与えずにプレイキャラクターの体力が 0 となるような状況が表現できなかったことによると分析されている。

なお、GA ではその特性上、遺伝子情報の初期値によって結果が左右される可能性に注意する必要がある。今回の学習においては、遺伝子情報の初期値は範囲内の一様乱数としていた。遺伝子情報の初期値をランダムとした条件の中で複数回学習を試したところ、いずれの学習結果においても、強さについてはおおむね同様の結果が得られている。

6.2 行動選択ログの分析による評価

今回、島モデル GA に採用した遺伝子パラメータは、NPC の行動を形成する基本情報ではあるが、実際に表出する行動とは必ずしも 1 対 1 に対応するものではない。そこで、学習の目標値を 6 段階のレベルで設定し、それぞれの目標値に対して島モデル GA を 1 回ずつ実施して得られた NPC 合計 600 個体を対象に、NPC どうしてゲームをシミュレートし、得られる行動選択ログに対しての分析を実施する。この際、対戦相手となる敵キャラクターについては、島モデル GA により得られた最も強い NPC とする。さらに、一定の確率でランダムな行動を選択する処理を付加することにより、対戦状況の多様性を確保する。

表 2 シミュレーションに使用した敵キャラクターのパラメータセット

Table 2 Enemy character's parameter set used for simulation.

パラメータ名	数値
$flame_{max}$	81
$W_{E_{hp}}$	36
$state_1$	28
$state_2$	35
$state_3$	18
$judge_1$	16
$judge_2$	42
$judge_3$	14

表 3 行動傾向分析のための分析対象

Table 3 Behavior selection log acquired at simulation.

番号	分析対象	単位
1	プレイキャラクターの残り体力	%
2	敵キャラクターの残り体力	%
3	近接攻撃回数	回
4	近接攻撃 Hit 回数	回
5	遠距離攻撃回数	回
6	遠距離攻撃 Hit 回数	回
7	防御行動回数	回
8	防御成功回数	回
9	敵からの近接攻撃 Hit 回数	回
10	敵からの遠距離攻撃 Hit 回数	回
11	歩き接近時間	フレーム数
12	走り接近時間	フレーム数
13	停止時間	フレーム数
14	歩き逃避時間	フレーム数
15	走り逃避時間	フレーム数

対戦シミュレーションで用いた NPC のパラメータセットを表 2 に記す。実験の評価対象は、前述の 6 段階のレベルを目標値として獲得したそれぞれのレベルで 100 個体ずつ、合計 600 個体とする。これら 600 個体に対し、1 回ずつのシミュレーションを行い、表 3 に記載した行動のログを取得する。表中、「プレイキャラクターの残り体力」および「敵キャラクターの残り体力」はシミュレーションの終了時点でのものである。「近接攻撃回数」から「防御成功回数」までは、ゲーム一試行中にプレイキャラクターが実施した行動選択の回数である。「敵からの近接攻撃 Hit 数」および「敵からの遠距離攻撃 Hit 数」は、対象が防御したか否かにかかわらず、攻撃を受けるたびにカウントしたものである。「歩き接近時間」から「走り逃避時間」までは、対象のそれぞれの行動時間の総計をフレーム数のカウント値として集計したものである。

以上のシミュレーションにより、各目標レベルごとに生き残った個体の行動選択状況のデータが取得されるが、個体が形成するクラスタの構想パターンの特徴をとらえるために、獲得した 600 個体すべてのデータを統合して GMM を用いた分析を実施する。なお、GMM に与えるクラスタ

数については、バイズ情報量規準（以下 BIC 値）を用いた分析による最適クラス数で与える。

6.2.1 行動選択ログの分析結果

上記手続きにより、各個体の行動傾向は、表 3 に記載した 15 次元のベクトルとして実測されるが、BIC 値により最適クラス数の分析をしたところ、6 つのクラスタを持つことが指示された。それぞれのクラスタ重心のプロットを図 5 に示す。

続いて、得られた 6 つのクラスタ間のクラスタ重心について、数学的な視点で異なっていることを確認するため、相関係数を用いて評価する。ここでは、各クラスタごとの組合せについて、間隔尺度を求めるため、ピアソンの積立相関分析を行う。6 つのクラスタ間について求めた相関係数を表 4 に示す。本調査では、行動選択ログにおけるクラスタ重心の相関係数を求めているため、負の相関を持つ組合せは、それぞれが相反する異なった行動選択傾向を持つことを意味する。正の相関を持つ組合せに対しては、帰無仮説を「無相関である」として検定を行ったところ、有意水準 5% で相関があるとされるクラスタの組合せは (No1, No3), (No1, No6), (No3, No6) の 3 つの組合せとなった。この結果は、クラスタの組合せ 15 件中、正の相関があるとされた 3 件を除いた 12 件が、数学的な視点において異なるものと判断されることを示唆している。

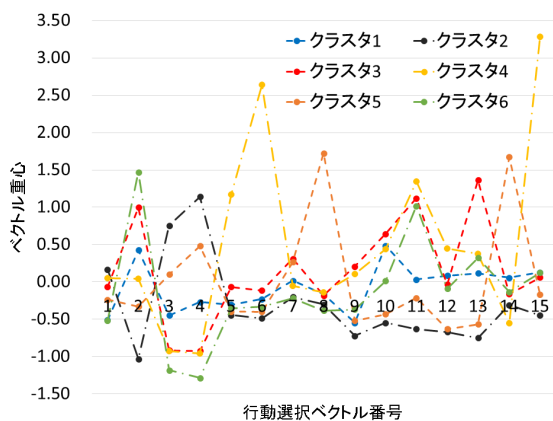


図 5 GMM により獲得された 6 つのクラスタの重心ベクトル
Fig. 5 The centroid vectors of the six clusters obtained by GMM.

表 4 クラスタ間の相関係数
Table 4 Coupling coefficient between clusters.

クラスタ番号	1	2	3	4	5	6
1	1.00	-	-	-	-	-
2	-0.52	1.00	-	-	-	-
3	0.61	-0.79	1.00	-	-	-
4	0.18	-0.46	0.26	1.00	-	-
5	-0.06	0.32	-0.39	-0.40	1.0	-
6	0.69	-0.81	0.87	0.34	-0.26	1.00

6.3 NPC の振舞いに関する被験者実験による主観評価

本節では、前節の実験で得られた 6 つのクラスタの特徴の差異を主観評価実験によって検証する。

6.3.1 実験手続き

上述の分析により典型的な行動パターンとして 6 つクラスタが形成されることが確認されている。ここでは、6 つのクラスタのそれぞれに属する最も強い NPC に対して、被験者に、1) 行動選択ログの取得の際に用いた NPC との対戦動画の視聴、2) NPC との対戦、を実施させ、それぞれの NPC に対する印象を以下のフォーマットで記述させた。

- (1) 設問 1：動画中のプレイキャラクタはどのようなプレイヤに見えましたか？
- (2) 設問 2：その理由をお答えください。

被験者は 22 歳から 24 歳のアクションゲームの経験がある男女 10 名であり、印象評価に先立ち、対象となる 2D アクションゲームを自由にプレイさせ、ゲームのルールや仕様を十分把握させた。NPC との対戦の際において、操作に戸惑った様子は見受けられなかった。

6.3.2 結果

実験結果は「被験者が各クラスタを代表する NPC に対しておおむね共通の印象をいただく」ことを示唆するものであった。設問 1 に対する回答例を表 5 に示す。

クラスタ 1 とクラスタ 3 の NPC は、強さの目標値を 20 としたときに生き残った個体であるが、クラスタ 1 に対しては、「消極的」といった評価が多く、「様子をうかがっているように見える」というコメントもあった。理由としては、「敵キャラクタの隙をうかがうために防御行動をしている」「相手の出方をうかがっている」といった事項があげられていた。クラスタ 3 については、「積極的」「バランス型」と評され、「近接攻撃・遠距離攻撃・防御行動がほどよく使われている」との理由があげられていた。強さの目標値が同じだった個体としては、上記のものほかに、目標

表 5 アンケート結果

Table 5 Questionnaire results.

クラスタ No	コメント
1	消極的, 不慣れ, 気まぐれ, 様子をうかがうプレイヤ 無謀, 焦っている, 単調, ふざけている, 初心者
2	憶病, 慎重, カウンタ型, 保守的なプレイヤ 面倒くさがり, 冷静, 敵前逃亡
3	積極的, バランスのとれたプレイヤ 中級者, 熟練者, 堅実, 試行錯誤している
4	遠距離攻撃型, 安全志向なプレイヤ せこい, 性格が悪そう, 憶病, 好まれない, 無知
5	近接攻撃型, 好戦的, 積極的なプレイヤ 先行逃げ切り, 直感型, 熟練者, 頭が悪そう ごり押しな性格, 攻撃性が強い, 捨て身
6	消極的, 逃亡型, 諦め思考, 適当な性格 遊んでいる, 勝つことが目的ではないプレイヤ

値を -20 とした際に得られたクラスタ 4 とクラスタ 6 があげられる。クラスタ 4 とクラスタ 6 に対しては、それぞれ「安全志向」「消極的」という、似通った「印象」があげられているが、その一方で、「遠距離攻撃型」「勝つことが目的ではない」というように、勝負に対するこだわりの部分での差がとらえられている様子がうかがわれる。

各クラスタの中で、最も積極的、好戦的と解されたのは、クラスタ 5 の NPC であった。この NPC は、強さの目標値を 100 とした際に生き残ったものであり、対戦動画の中では、唯一プレイキャラクター側が勝つことができるものであった。

7. 考察

本研究は、先験的な情報を与えずに NPC の多様な振舞いを構成することを目標とし、島モデル GA を用いてゲーム AI の導出、強さのレベル制御の可能性と行動傾向の分析を実施した。行動選択ログを用いたクラスタリングによる評価からは、学習の目標値ごとに複数の行動傾向を持つクラスタが獲得されていることを示す結果が得られた。また、獲得されたゲーム AI の振舞いについて主観評価実験を行ったところ、観察者が、それぞれの AI に対して、ほぼ共通した印象をいただいている様子を示す結果が得られた。「積極的」や「消極的」、「遠距離攻撃型」など行動自体をい表す評価語に加えて、「慎重」「気まぐれ」「保守的」など、より感覚的な評価語、さらには、「せこい」や「性格が悪そう」といった、NPC 自体のパーソナリティを表すような評価語も使われた。

これらのうち、クラスタ 1, 3 については、 f_{flame} の値がともに高いにもかかわらず、「焦っているプレイヤー」「バランスのとれたプレイヤー」と相反する評価がなされていることが分かる。これは、 f_{flame} の値により、どちらの NPC も、比較的長時間連続して操作が行われているが、 $state_1$ から $judge_3$ までのパラメータ設定の違いにより、攻撃行動や防御行動が適切なタイミングで行われているかいないかの違いであると考えられる。また、NPC どちらのシミュレーション結果を見ると、クラスタ 2, 5 において近接攻撃行動が他のクラスタよりも多いことが分かる。これは、図 2 における条件判定 C_1 で用いられるパラメータ $judge_1$ の値が他よりも極端に低いことが影響していると考えられる。以上のように、NPC の実際の行動や、人間が受ける印象に対して、今回島モデル GA により最適化を行った BehaviorTree の各パラメータが影響していることが分かった。

本実験においては、「積極的」「消極的」、あるいは、「性格が悪そう」という概念に対応した行動パターンに関する先験的な知識をいっさい与えていない。その状況で、人間が個性や性格を見出す NPC の行動パターンを獲得できた。このことは、ゲーム制作における省力化という目的に

おいて重要な貢献の 1 つとなりうると思う。また、本研究での実験プロセスおよびその結果は、ゲームを題材として、概念の形成のシミュレーションを実施したという側面も有している。今後は、双方の視点で、研究を進展させていきたい。前者の視点に関しては、「好戦的」と評価されたゲーム AI に代表されるように、学習の目標値自体が行動パターン、ひいては、その印象に大きな影響を与えることも確認されている。今後は、この部分のより精緻な検討を行うとともに、他のゲームジャンルについても実験を実施していきたい。一方、後者については、現時点では、ごく限られた状況での実験を実施できたにすぎない。敵キャラクターのパターンを増やした状況での実験のほか、大規模な評価実験を実施していく必要がある。

8. おわりに

ゲームデザインにおいて、NPC のキャラクターの設定、および、振舞いのデザインは、経験豊かなデザイナーの作業によって実践されてきた。本論文では、キャラクターとは、与えられた制約条件下で最適化が行われ典型化した振舞いに対して付与されたラベルであるという考えのもと、島モデル GA を用い、先験的な情報を与えずに NPC の振舞いのクラスタを獲得し、そのクラスタの特徴を、数学的アプローチ、主観評価の評価により分析を試みた。

このアプローチにより、独自に開発した 2D アクションゲームに対して複数の強さ目標を設定して NPC の BehaviorTree におけるパラメータの最適化を進めた結果、典型的な振舞いを有する 6 種類の NPC が選出され、人間の評価者は、そのそれぞれに対して、「積極的」「消極的」「焦っている」「保守的」など、主観に相当する用語によって参照されうる状況が確認された。

本研究は、EC 領域の文脈でいえば振舞いのデザインの自動化、さらには、概念形成に関する構成論的アプローチの一研究としての意義を有していると思う。本論文では、2D アクションゲームと対象とした研究についての報告を行ったが、対象を広げてさらなる実験を進めていきたい。

参考文献

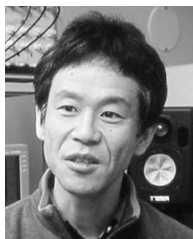
- [1] 一般社団法人情報処理学会：コンピュータ将棋プロジェクト、入手先 (<https://www.ipsj.or.jp/50anv/shogi/index2.html>) (参照 2016-12-13)。
- [2] Benbassat, A. and Sipper, M.: Evolving lose-checkers players using genetic programming, *Proc. 2010 IEEE Conference on Computational Intelligence and Games*, IEEE, pp.30-37 (2010)。
- [3] 藤井叙人, 佐藤祐一, 若間弘典, 風井浩志, 片寄晴弘: 生物学的制約の導入によるビデオゲームエージェントの「人間らしい」振舞いの自動獲得, *情報処理学会論文誌*, Vol.55, No.7, pp.1655-1664 (2014)。
- [4] 上田陽平, 池田 心ほか: 遺伝的アルゴリズムによる人間のレベルに適應する多様なオセロ AI の生成, 第 27 回ゲーム情報学研究会 (2012)。

- [5] 星野准一, 田中彰人, 濱名克季: 模倣学習により成長する格闘ゲームキャラクター, 情報処理学会論文誌, Vol.49, No.7, pp.2539–2548 (2008).
- [6] Fujita, H. and Ishii, S.: Model-based reinforcement learning for partially observable games with sampling-based state estimation, *Neural Computation*, Vol.19, No.11, pp.3051–3087 (2007).
- [7] Hoki, K. and Kaneko, T.: The global landscape of objective functions for the optimization of shogi piece values with a game-tree search, *Advances in Computer Games*, pp.184–195, Springer (2011).
- [8] Togelius, J., Karakovskiy, S. and Baumgarten, R.: The 2009 Mario AI competition, *IEEE Congress on Evolutionary Computation*, pp.1–8, IEEE (2010).
- [9] Pepels, T., Winands, M.H. and Lanctot, M.: Real-Time Monte Carlo Tree Search in Ms Pac-Man, *IEEE Trans. Computational Intelligence and AI in Games*, Vol.6, No.3, pp.245–257 (2014).
- [10] Esparcia-Alcázar, A.I., Martínez-García, A., Mora, A., Merelo, J. and García-Sánchez, P.: Controlling bots in a first person shooter game using genetic algorithms, *IEEE Congress on Evolutionary Computation*, pp.1–8, IEEE (2010).
- [11] 佐藤 浩, 小野 功, 小林重信: 遺伝的アルゴリズムにおける世代交代モデルの提案と評価, 人工知能学会誌, Vol.12, No.5, pp.734–744 (1997).
- [12] Eshelman, L.J.: Real-Coded Genetic Algorithms and Interval-Schemata, *Foundations of Genetic Algorithms*, Vol.2, pp.187–202 (1993).



福嶋 良平 (学生会員)

1993年生。2015年関西学院大学理工学部人間システム工学科卒業。2017年同大学大学院理工学研究科人間システム工学専攻修士課程修了。主にゲームAIの研究に従事。



片寄 晴弘 (正会員)

1991年大阪大学大学院基礎工学研究科博士課程修了。工学博士。イメージ情報科学研究所, 和歌山大学を経て, 現在, 関西学院大学理工学部教授。エンタテインメントコンピューティング, 音楽情報処理, 感性情報処理の研究に従事。

科学技術振興機構さきがけ研究21「協調と制御」領域研究者。科学技術振興機構CREST「デジタルメディア(略称)」領域CrestMuseプロジェクト研究代表者。電子情報通信学会, 人工知能学会各会員。