

感情演出による楽しませる対戦型格闘ゲーム AI

南 基大^{1,a)} 池田 心^{1,b)}

概要: 人間のゲームプレイを観察していると、「調子に乗って強気な行動をしてしまう」ことや、「やる気を損ない、投げやりな行動を取る」といった感情に起因した行動が確認できる。これらの行動傾向の変化は「多様性」や「人間らしさ」を演出すると推測され、ゲーム AI に取り入れることでゲームの楽しさ向上が期待できる。しかし、このような行動傾向の変化は「高スコアを取る」や「対戦相手を倒す」などのゲームの主たる目的のみを達成しようとするゲーム AI には生まれにくい行動である。そこで我々はゲーム AI に感情に起因した行動を取り入れることで、ゲーム AI に対し人間らしさの演出を試みた。まず、被験者 12 名にゲーム AI と対戦してもらい、状況ごとの行動とそのときの感情を採取した。そして、状況と感情のペア群から、教師あり学習によって感情推定器を構築した。さらに、状況と行動のペア群を、感情によって分割したうえで教師あり学習することで、「喜び時行動モデル」等 4 つの感情行動モデルを作成した。教師あり学習に加えて、より明確な行動傾向の変化を演出するため、特殊な報酬を用いた強化学習による感情行動モデルも作成した。対戦評価実験の結果、「楽しさ」「行動傾向変化の自然さ」について提案手法が優れていること、強化学習によるモデルのほうがより良い結果となることが分かった。

キーワード: 対戦型格闘ゲーム, 人間らしさ, 感情, 楽しませる

Human-Like AI Players Considering Emotions in Fighting Games

MOTOHIRO MINAMI^{1,a)} KOKOLO IKEDA^{1,b)}

1. はじめに

近年の計算機の性能向上などにより、人工知能技術は著しい発展をみせている。その中でもゲーム AI 領域における活躍は著しい。ボードゲームのような比較的単純なゲームでは既にゲーム AI が人間のトッププロプレイヤーを上回っているが、より複雑で膨大な計算量を求められるビデオゲームにおいても人間以上の性能をもつゲーム AI が出つつあり、「強さ」に関するゲーム AI の研究はある程度達成されつつあるといえる。

ゲーム AI における研究分野として「強さ」の他に、「人間を楽しませる」ことを目的としたゲーム AI の研究があり、中でも「人間らしい」ゲーム AI によって人間を楽し

ませる研究も注目されている。Sila らは、状況に応じて人間の感情が変化することによってプレイ傾向にも変化が生じることを再現するために、複数の行動モデルを遷移させる仕組みの提案とマリオ風横スクロールアクションでの実装をしている [1]。残り時間が少ない状況での急ぎモデルや、コイン収集を重視した強欲モデル等の複数の行動モデルがあり、ゲーム情報を元に手書きのルールベースのスイッチによってこれらの行動モデルを切り替え、人間の感情と行動傾向変化を表現している。

以上のような、ゲームの状況からの感情を決定し、行動傾向を変化させるゲーム AI はここ数年でいくつかの研究がされている。これらの従来研究において、感情決定の多くは手書きのルールベースによって作成されている。しかし、ゲームの複雑さが増すと手書きのルールベースで全ての状態を記述するのが難しなると考える。そこで本研究では、人間プレイヤーのプレイログによる学習に過去の感情タグを合わせた、ラーニングベースのモデルによる感情決定

¹ 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology
a) s2010169@jaist.ac.jp
b) kokolo@jaist.ac.jp

モデルを提案する。この感情決定モデルを用いて、複数の感情的行動モデルを切り替える人間らしいゲーム AI の提案と実装を行う。そのために、感情収集アンケート方法の提案、人間による感情情報付きのプレイログの収集、教師有り学習による感情決定器の作成、教師あり学習及び強化学習による感情的な行動モデルの獲得を行う。

2. 感情とゲームについての関連研究

2.1 感情の種類と分類

人間感情の分類に関する研究は古くから行われているが、感情の数と種類を大まかに分類する「基本的感情 (basic emotions)」が近年の流れになっている。Ekman は人間の表情が感情に影響されることに着目し、基本的感情を「喜び」「恐れ」「驚き」「悲しみ」「嫌悪」「怒り」と6種類に分類することを提唱している [2]。この他にも基本的感情の分類には数多くの提唱がなされているが、多くの提唱で「喜び」「怒り」「悲しみ」「嫌悪」「恐れ」の5種類が共通しており、後述する関連研究においてもこの5種類の基本的感情がよく用いられている。本研究においても上記5種類を基本的感情として定義して利用することとした。

2.2 ゲームと感情の相互作用

第1章で述べたように、ゲーム AI に感情を用いて人間らしさを演出している研究はいくつか行われている。

飯田らは不特定多数の人のゲーム中における感情の調査と分析を行なっている [4]。ゲームにおいて自陣側が優勢、若しくは接戦であるとき、「興奮」や「快適」の感情となり、対戦相手が優勢である時は「いらだち」を感じる傾向を示している。木村らは「素早い反応を要求するデジタルゲームは人間プレイヤーの感情に特有の変化を生じさせる」ことを提唱し、検証実験を行っている [3]。素早い反応を求められる対戦格闘ゲームと、反応を求められないシュミレーションゲームのプレイヤー感情を比較検証し、素早い反応が求められるゲームでは「喜び」や「怒り」などの興奮度の高い感情が比較的多く想起させられることを示している。

Sila らは、ゲームの状況からプレイヤーの感情を決定し、コイン収集やタイムアタックなどの副目標を目的とした「焦り傾向」や「強欲傾向」等の複数の行動モデルを切り替える仕組みを提案している [1]。このような複数の行動モデルを用意し切り替える手法は行動モデル獲得の容易さに優れており、本研究の提案手法においても同様の手法を利用する。

3. 提案手法

通常人間がゲームを遊ぶ際は「敵を倒す」や「高スコアを取る」等のゲームの主目的の達成を目指し、そのために自身が最適と考える行動を選択することが多い。しかし、人間がゲームを遊ぶ様子を観察していると、ゲーム中の行

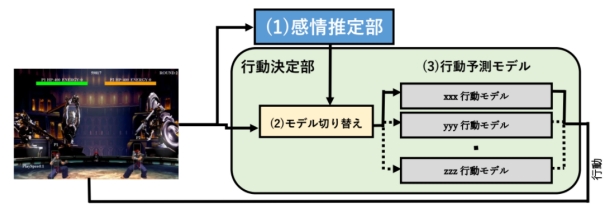


図 1 提案手法フレームワーク

動傾向に様々な変化が現れる事がある。例えば、人間プレイヤーは一方的に負け続けることで戦意や集中力を損ない、弱腰な動きを取ることがある。これらは「ティルト」や「萎え」と呼ばれている。このように人間プレイヤーの感情がゲームの状況によって変化し、行動選択にも影響を与えることは一般的に知られている。

我々は感情をゲーム AI に導入するために、「感情決定部」と「行動決定部」の2層によって構築されたフレームワークを提案する (図 1)。また、環境にはゲーム AI の構築が比較的容易で、感情が重要な要素として知られている対戦型格闘ゲームを利用する。

3.1 感情決定部

感情決定部では、ゲームの情報から人間であればどのような感情となるかを決定し、決定した感情情報によってゲーム AI の行動モデルの切り替えを行う。

先行研究における感情決定部の多くはルールベースモデルによって作成されている。ルールベースモデルは設計者が if-then 文等によりゲーム内の状況に応じて感情を記述することで実現される手法となる。しかし、設計者がゲームのルールや文化、プレイヤー心理などを詳細に把握する必要があり、ゲームの自由度や複雑性によっては表現が不足すると考える。本研究では対戦型格闘ゲームを環境に用いるが、対戦型格闘ゲームはキャラクターの位置や行動の選択肢、体力残量などの組み合わせにより膨大な量の状況が存在する。さらに過去の感情と現在の感情も組み合わせる必要がある。ルールベースでは全ての状況における人間感情を記述するのは限界があると考えられる。

そこで本研究では感情情報がタグ付けされた人間のプレイログによる教師有り学習を用いた、ラーニングベースの感情決定モデルを提案する。ルールベースでは網羅しきれない範囲であっても、十分な量のデータと表現力の高いモデルであれば人間の感情をある程度は模倣可能と考える。

また、収集した感情タグ付きプレイログを用いて感情決定部の構築を行うが、我々是对戦型格闘ゲームの“時系列性”に注目した。対戦型格闘ゲームのある状況から感情を予測する際、その瞬間にゲームを観測して得られる情報はその時点での「体力差」「経過秒数」「キャラクタ位置」などになる。しかし、対戦型格闘ゲームはプレイヤーの過去の感情状態や行動傾向の癖など、その瞬間のゲーム情報では

観測できない情報がプレイヤーの感情に大きな影響を与える
と考える。そこで感情決定部に教師あり学習のネットワー
ク構造に時系列情報を考慮できる構造を導入し、感情決定
の精度向上を狙う。

3.2 行動決定部

本節では、ゲーム中でのキャラクター操作を決定する
「行動決定部」について説明する。

行動決定部には、「ニュートラル」「喜び」などの各感情
状態に紐付けられた複数の行動モデルが含まれており、感
情決定部で得た感情情報を元に行動モデルの切り替え・推
論を行うことで、人間の感情に影響された行動を模倣する。
これら各感情に紐付けられた行動モデルを、本研究におい
て「感情的行動モデル」と呼称する。また、本研究では人
間のプレイログから学習する教師あり学習モデル（以下 SL
モデルと呼称）と、報酬設定により意図的に傾向付けられ
た強化学習モデル（以下 RL モデルと呼称）の 2 種類を構
築する。

SL モデルは後述するプレイログ収集実験で獲得したデー
タセットから感情タグの種類ごとにデータを分類、学習に
用いることで人間の各感情別の感情的行動モデルの獲得を
狙う。実際の人間によるプレイログを用いることで、より
現実に近い形で感情的行動モデルの獲得が見込める。

しかし、感情による行動の変化をゲーム AI が獲得でき
たとしても、その差異を人間が認識することも出来ない事
も想定される。そこで各感情毎の人間の行動を分析し、強
化学習によって意図的な変化を加えた行動モデルの獲得を
目指す。各基本的感情ごとに報酬を設定し、ニュートラル
行動モデルと感情的行動モデルに明確な差を設ける。

4. 研究環境

4.1 対戦型格闘ゲーム

対戦型格闘ゲームは主にコンピュータ上で遊ばれる 2 人
対戦ゲームである。各プレイヤーによって操作されるキャラ
クターが一定範囲のステージ内で互いに格闘技や遠距離攻撃
を用いて相手の体力を減らし、制限時間内に相手の体力を
全てなくすか、制限時間時に体力が多いほうが勝利となる。

対戦型格闘ゲームは非常に素早くゲーム状況が推移して
いくことが知られている。そのため、相手の攻撃を見てか
ら対応することは難しく、「相手は焦っているから必殺技
を打ってくるはず」や「相手は調子にのってジャンプする
から対応しよう」などの相手の感情や行動などを予測して
行動することが重要となり、これは「読み合い」として一
般的に知られている。このように対戦型格闘ゲームは他対
戦型ゲームと比べて感情の重要度が高く、感情によって人
間らしさを演出する際に有利であると考えられる。そこで本
研究では実験環境に対戦型格闘ゲームを利用する。



図 2 FightingICE の対戦画面

4.2 FightingICE

FightingICE は、Java で構築された対戦型格闘ゲームの
AI 研究用オープンソースプラットフォームである [5]。

本研究はゲーム AI の「強さ」ではなく「人間らしさ」の向
上为目标であるが、「FightingICE 及び国際大会参加 AI が
オープンソースで公開されている」「OpenAI GymAPI が
公開されており、豊富な機械学習ライブラリを持つ Python
が利用できる [6]」「人間用のキーボードインターフェイス
が用意されている為、被験者実験が容易」等の理由から
FightingICE を環境として利用する。

また、FightingICE は主に強さを求めたゲーム AI の学
習環境向けに開発されている。そのため、人間がプレイす
る際には快適性などに幾つかの問題があった。そこで本研
究では FightingICE に対して、「操作の簡易化」や「攻撃
技種類の削減」等、独自に 20 箇所以上の改修を施した。改
修には人間向けに調整されている市販ゲームを参考にし、
対戦型格闘ゲームで人間が本来生じるはずの感情が妨害さ
れないよう配慮した。

5. 感情タグ付きプレイログ収集実験

3.1 節で述べた感情決定部の学習には、人間の感情がタグ
づけされたプレイログが必要となるが、そのようなデータ
セットは公開されていない。そのため、対戦型格闘ゲーム
におけるプレイヤーのプレイログと感情、ゲームデータ等の
収集を行う必要がある。以降本節では、対戦型格闘ゲーム
における感情タグ付きプレイログ収集について説明する。

5.1 感情測定

対戦型格闘ゲームのような素早い進行速度のゲームにお
いて、感情は刻々と変化することが予測される。プレイヤー
の操作を阻害するような感情測定のアンケートが挟まると、
アンケート回答中に感情が変わってしまい、正しいデータ
にならないことが懸念される。そこで本研究ではリプレイ
再生と感情採取アンケートをあわせて用いることで、ゲー
ムの流れを遮らず、かつ一定時間以内での感情収集を可能
とするデータ収集環境を構築することとした。

5.1.1 感情測定のタイミング

対戦型格闘ゲームでは、パンチなどの数分の1秒の攻撃を起点に大きなダメージを与えられるため、1つのミスで大きな劣勢につながるが多々ある。そのため、通常の戦闘場面ではプレイヤーは感情を意識せず、キャラクターの操作に集中することが多いのではないかと幾つかの主観的経験に基づき考える。つまり人間プレイヤーが感情を強く意識するのは、プレイヤーの操作密度が低い状況であると考えられる。

そこで我々は対戦型格闘ゲームにおける特徴的な状況である、「ダウン状態」に注目した。「ダウン状態」は、キャラクターが特定の技や一定以上のダメージを受けた際に倒れることで、「数秒」の間「対戦相手から攻撃」「プレイヤーの操作」を受け付けられない状態のことである。

そこで本実験では、感情測定を行うタイミングをリプレイ再生中の「キャラクターのダウン状態時」とすることにした。

5.1.2 感情測定手法

本研究では、比較的短時間でのアンケートを実施するため、独自のアンケートツールを作成した。このさいに、織田らの「感情・覚醒チェックリスト (Emotion and Arousal Checklist)」[8]を参考にした。この方法では、「喜び」「怒り」「悲しみ」「嫌悪」「恐れ」の計5種の感情に関連した20種の用語で四件法によって感情測定を行っている。この20用語の因子負荷量を参考に、対戦型格闘ゲームに適しているであろう単語を5種の基本的感情ごとに1用語、合計5種の用語を選出し、これを感情語として用いる事とした。各感情と用語の対応は以下のようになる。

- 「喜び」→「楽しい」
- 「怒り」→「いらいらする」
- 「悲しみ」→「落ち込んでいる」
- 「嫌悪」→「不快な」
- 「恐れ」→「こわい」

アンケート実施時に、回答者には「プレイ中の感情を思い出したとき、最も当てはまる用語はどれですか?」と質問した。回答の選択肢には感情語5語と強度3種の合計15択を用意し、最も当てはまる感情と強度の組み合わせを1つを選択してもらった。

5.2 実験方法

実験に参加した被験者は大学生13名(男性12名女性1名、20代から30代)である。各被験者は、事前に指定された対戦順に従い、対戦とリプレイ視聴及び感情調査アンケート回答を繰り返し行ってもらった。対戦相手には6種のゲームAIを用い、合計60試合の対戦を実施した。実験手順を以下に説明する。(1)被験者は指定されたゲームAIと対戦を行う。被験者はゲームパッドを用いてキャラクターを操作し、敵キャラクターに対して勝利を目指すよう指示されている。1試合1ラウンド制、1ラウンドの制限時

間は60秒に設定した。(2)1試合終了後、開始ボタンを押すことでリプレイが再生される。被験者には「プレイ時の感情を思い出しながら視聴してください」と指示した。リプレイ再生中、キャラクターがダウンするか一定時間が経過した際に画面が一時停止し、その時点での感情に関するアンケートに回答が指示される。被験者がアンケートに回答すると自動的にリプレイ再生が再開される。被験者には再生終了まで視聴とアンケート回答を繰り返してもらった。(3)対戦順番が記載された用紙に該当番目の完了チェックを記入し、実験終了まで(1)と(2)を繰り返す。

また、被験者には実験終了後に、5種の基本的感情が発生しやすかった状況を回答してもらった。

5.3 収集されたデータの統計量

本実験では、合計で650試合分のデータを収集した。また、得られた感情情報は合計で9654個であり、感情採取間隔の平均は3.70秒だった。表1に強度別の5種の基本的感情の発生および自覚の頻度を示す。5種の基本的感情の分布は、「喜び」が52%、「怒り」が34%、「嫌悪」が8%、「悲しみ」が5%、「恐れ」が1%であった。「喜び」と「怒り」の2種の基本的感情で全体の86%と大きく、「恐れ」は全体の1%と僅かだった。

本研究では、これらの感情の生起理由を考察したうえで、恐怖に関しては再現の対象から外し、比較的近い悲しみと嫌悪については一つのグループにまとめることにした。すなわち、感情決定器は「喜び」「怒り」「悲しみ・嫌悪」の3種を決定し、感情行動モデルとしてはそこに「ニュートラル」を加えた4種の行動モデルを作成することにする。

	喜び	怒り	悲しみ	嫌悪	恐怖
	楽しい	いらいら	落ち込む	不快	怖い
強	1706	512	116	226	10
中	2405	1647	263	315	61
弱	950	1112	118	185	28

表1 5種の基本的感情及び強度別分布

6. 感情決定器

感情決定部は、キャラクターの体力や位置、使用技などをゲーム状況を保持し、人間プレイヤーだとしたら持つだろう感情の予測を行う。本項では、感情決定器の機能を改めて説明したのち、実装に用いた各種データや、予測性能について述べる。

6.1 感情決定アルゴリズム

3.1節で述べたように、感情決定アルゴリズムには時系列分類手法のLSTMを用いる。本実験では(1)全結合層(2)LSTM層(3)出力層の3層構造を用いた。表2に学習パラメータを示す。

パラメータ	値
入力次元数	153
全結合層次元数	300
活性化関数	ReLU
LSTM ブロック数	10
出力次元数	3
epoch	64
batch_size	32
最適化手法	adam
誤差関数	categorical_crossentropy

表 2 感情決定モデル各種パラメータ

6.2 学習用データセット

学習に用いたデータ数は合計で 4314 セットであり、内約は「喜び」が 1944 セット、「怒り」が 1671 セット、「悲しみ/嫌悪」が 699 セットとなる。学習を行う際の入力には、「過去 10 秒分におけるゲーム状況」、「過去 3 回分の感情決定値」の 2 種類を用いている。また、出力は、「喜び」「怒り」「悲しみ/嫌悪」の 3 種のクラス分類となる。

6.3 性能評価

獲得した感情決定モデルに対し、テスト用データに 120 試合 1729 セット分を用いて性能評価を行った。学習用データとテスト用データは別の試合から抽出している。また、テスト用データの感情タグ内約は、「喜び」648 セット、「怒り」557 セット、「悲しみ/嫌悪」524 セットとなる。

結果として、感情決定モデルはテストデータに対し、正解率が 0.62 を示した。また、混同行列を図 4 に示す。

「喜び」について比較的正しく予測できているのに対し、「怒り」「悲しみ/嫌悪」については誤った予測がある程度みられた。これには、「喜び」は訓練データが多い点、「怒り」「悲しみ/嫌悪」についてはネガティブな感情という意味で似ている点などの可能性があると考えている。

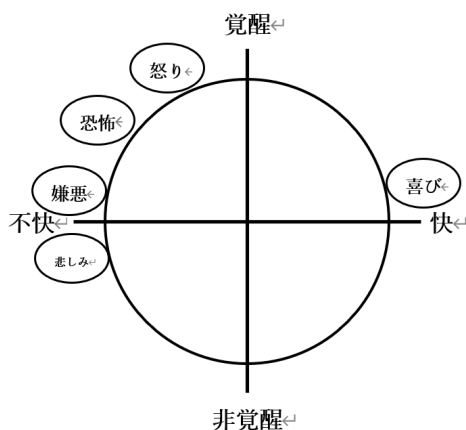


図 3 russell の感情モデルによる基本的感情の表現 (Russell を参考に作成)

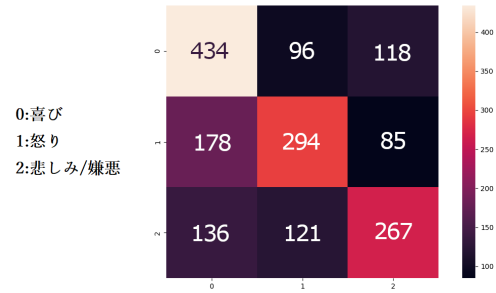


図 4 感情決定部評価実験における混同行列

7. 行動決定部

行動決定部は、各感情の影響を再現した複数の感情的行動モデルで構成される。「喜び」「怒り」「悲しみ/嫌悪」の 3 種の基本的感情に紐付けられた 3 種類の行動モデルに、初期状態に用いるニュートラル行動モデルを加えた計 4 種である。

また、感情的行動モデルの獲得は、教師あり学習と強化学習の 2 つの方法で行った。本項では、学習方法別に、実装に用いた各種データやパラメータ、性能について述べる。

7.1 教師有り学習による実装

教師あり学習による感情的行動モデル (以降、SL モデル) では、感情タグ付きプレイログ収集実験で得たゲーム状況とプレイヤーの操作ログのペアを用いて学習することで、可能な限り人間の行動に近づけることを意図している。

7.1.1 行動モデルのアルゴリズム

各行動モデルのアルゴリズムには、ニューラルネットワークによる分類手法を用いる。ニューラルネットワークの構成は (1) 全結合層 (2) 全結合層 (3) 出力層 の 3 層構造を用いた。モデルの説明変数は 144 次元のゲーム状況であり、目的変数は 21 次元のキャラクタ操作となる。

表 3 に各種パラメータを示す。

パラメータ	値
入力次元数	144
全結合層 (1 層目) 次元数	300
全結合層 (2 層目) 次元数	300
活性化関数	ReLU
出力次元数	21
epoch	128
batch_size	32
learning late	0.0001
最適化手法	adam
誤差関数	categorical_crossentropy

表 3 SL モデル各種パラメータ

7.1.2 学習用データセット

「ニュートラル」「喜び」「怒り」「悲しみ/嫌悪」の4種類の感情的行動モデルを作成するため、学習データについても4種類の作成を行った。「喜び」「怒り」「悲しみ/嫌悪」の3種類の感情的行動モデルについて、それぞれの感情タグ毎から、90,000frameをランダムに抽出し、各感情別学習用データとした。学習データの説明変数には、5章の実験で収集した各時点 t でのゲーム状況 s_t とプレイヤー操作 A_t のペア (s_t, a_t) を利用する。学習データの目的変数はキャラクター行動21種の離散値を利用したクラス分類モデルとなる。

7.1.3 性能評価

獲得した各行動モデルを評価するため、学習用データと同じ条件下で、4種類の感情ごとに10,000セットのテストデータを抽出した。各感情的行動モデルは対応するテストデータに対し、「ニュートラル」が0.38、「喜び」が0.33、「怒り」が0.32、「悲しみ/嫌悪」が0.29の一致率を示した。

SLモデルの各感情的行動モデルの強さを確認するため、FightingICEに標準搭載されており、本環境では人間プレイヤーの初心者相当の強さである「MctsAI」に対しSLモデルで100試合の対戦を行った。結果として、それぞれの勝率が「喜び」が64%、「怒り」が61%、「悲しみ/嫌悪」が53%となった。また、感情決定部と組み合わせを行ったSLモデルは、63%の勝率となった。モデルごとに差があるが、強すぎず弱すぎない人間初心者程度の強さを獲得できたと考える。

7.2 強化学習による実装

著者らの経験からも本研究で行ったインタビューからも、人間のプレイが感情によって左右されることは確かだと考えるが、それがリアルタイムで対戦している相手に伝わるかどうかはまた別問題である。たとえSLモデルで人間の感情行動が十分模倣できたとしても、それでは「人間らしい」だけであり、「人間らしく見える」ためには不十分である可能性もある。

そこで本研究では、意図的に感情の演出が強調された行動モデルが必要と考え、強化学習による行動モデル（以降、RLモデル）の構築を行った。ただし、ニュートラル行動モデルに関しては、SLモデルのものを流用している。

RLモデルの強化学習アルゴリズムには、深層強化学習手法のDQN(Deep Q-Learning)を用いる[7]。また、ニューラルネットワーク部に関しては、教師あり学習と同じパラメータを用いており、構成は(1)全結合層(2)全結合層(3)出力層の3層構造を用いた。

7.2.1 感情ごとの行動傾向と報酬設定

RLモデルでは、感情の演出が強調された行動モデルを構築することで、人間プレイヤーが「行動の変化」と「各感情おける動きの癖」を認識しやすくなることを狙っている。

そのため、各感情的行動モデルでは、それぞれの行動傾向の差が明確であること、各行動傾向の意図をプレイヤーが認識できることの2点が求められる。

そこで、対戦型対戦ゲームの経験者に感情に関する印象を調査し、それに基づき「喜び」「怒り」「悲しみ/嫌悪」の3種類の感情に対し、それぞれの行動傾向を設定した。次に、各行動傾向に沿った特殊な報酬設計を行うことで、各感情の行動モデルを獲得した。行動傾向及び報酬はそれぞれ複数種類を準備したが、その中でも学習及び感情的行動の演出が成功したものを筆者の主観的評価で選択している。

また、DQNは報酬を $[-1, 0, 1]$ に固定する'clipping'を行う事が多いが、行動ごとに報酬の差をつけるために今回の実験ではこれを用いないこととした。

以下に、「喜び」「怒り」「悲しみ/嫌悪」の3種類の感情別に設定した行動傾向と報酬設計を述べる。

・喜び→強気な行動

「喜び」感情は強気な行動傾向とし、以下のような報酬設定を行なった。

- ダッシュやジャンプなどで距離を詰める行動に対して微弱な報酬を与えた。
- キャラクター間の距離が近いとき、対戦相手に与えたダメージに応じた報酬を与えた。対戦相手から受けたダメージは考慮しないものとした。これにより、相手からのダメージを気にせずとにかく距離を詰め攻撃する強気な行動を獲得させることを狙った。

・怒り→単調化

「怒り」感情は単調な行動傾向とし、以下のような報酬設定を行なった。

- 対戦相手に与えたダメージに比例する報酬を与えた。また、対戦相手から受けたダメージに応じた負の報酬を与えた。
- 最も発動が早い技と最もダメージの大きい技の2種でのみ、対戦相手に与えたダメージに応じた報酬を与え、その他の攻撃技によるダメージでは報酬を与えないものとした。これにより、この2種の攻撃のみの単調な挙動を獲得させることを狙った。

・悲しみ/嫌悪→弱腰

「悲しみ/嫌悪」感情は弱腰な行動傾向とし、以下のような報酬設定を行なった。

- 距離を詰める行動に対して微弱な負の報酬を与えた。
- キャラクター間の距離が近いとき、距離を離す行動に対して微弱な報酬を与えた。
- キャラクター間の距離が遠いとき、射程の長い技でのみ、対戦相手に与えたダメージに応じた報酬を与え、そのほかの技では報酬を与えないものとした。これにより、とにかく距離をとり、消極的に遠距離攻撃を行う挙動を獲得させることを狙った。

7.2.2 性能評価

RL モデルの各感情的行動モデルの強さを確認するため、FightingICE に標準搭載されている ‘MctsAI’ と各感情行動モデルで、それぞれ 100 試合の対戦を行った。結果として、「強気な行動モデル」が 65%、「単調化行動モデル」が 46%、「逃げ腰行動モデル」が 43%となった。各感情行動モデルごとに強さの差はあるが、対戦していて楽しめる強さに収まっていると考える。

7.3 感情的行動モデルの切り替え

行動決定部による行動選択時には、感情決定部から受け取った感情情報を用い、「喜び」「怒り」「悲しみ/嫌悪」の各感情的行動モデルを切り替えて行動の選択を行っている。また、感情決定部の都合上、試合開始から最低 10 秒間は感情情報の更新が行われない。そのため、初回の感情情報の更新が行われるまでは、ニュートラル行動モデルを用いて行動選択を行う。

8. 評価実験

本節では、SL と RL の 2 種のモデルの有効性を検証するために実施した、動画視聴と対戦の 2 種類の評価実験について述べる。

8.1 実験内容

実験に参加した被験者は大学生 12 名（男性 12 名、全員 20 代）であり、公正な評価のために 5 章の実験とは異なる被験者のみを採用した。本評価実験は、動画視聴実験と対戦実験の 2 つの方法によって評価を行う。

動画視聴による評価実験では、構築した RL/SL モデル、ニュートラルモデル、既存の AI2 つ、および初級者*1 の計 6 プレイヤを、それぞれ別に用意した強さの異なる 3 つの AI と対戦させた。その後、その動画を被験者に視聴してもらった後、「人間らしさ」「行動傾向変化の自然さ」についてアンケートを実施した。この動画視聴実験に用いる動画は、6 種類のプレイヤと 3 種類の対戦相手で計 18 種の組み合わせである。各組み合わせ毎に 2 回の撮影を行っており、合計で 36 種の動画を用いている。視聴する動画は被験者全員に共通のものになるが、各被験者ごとに再生順番はランダムに入れ替えている。

対戦による評価実験では、構築した RL/SL モデル、ニュートラルモデル、既存の AI2 つ、の計 5 プレイヤと対戦してもらい、「楽しさ」「行動傾向変化の自然さ」「動きの自然さ」の 3 つの項目についてアンケートを実施した。対戦は 5 種類の AI に対し 3 回ずつ実施するため、被験者は合計で 15 回の対戦を行う。また、対戦する順番は被験者毎に

ランダムに入れ替えを行っている。

両実験とも、アンケートでは被験者に各項目ごとに 0 から 4 の 5 段階リッカート尺度によって回答してもらった。また、実験終了後に被験者には、「人間らしいと判断した基準」と「AIらしいと判断した基準」について、それぞれ 1 件を回答してもらった。

8.2 実験結果

本節では、各評価実験で得られた結果を示す。

8.2.1 動画視聴による評価実験

動画視聴実験における「行動傾向変化の自然さ」「動きの自然さ」について、6 種の評価対象別に平均値と標準偏差を表 4 に示す。

「行動の自然さ」「行動傾向変化の自然さ」の両項目において、提案手法は人間や既存 AI に劣る結果となった。

原因として、行動モデルの表現力不足が推測される。実験終了後の感想アンケートでは、被験者が「対象をゲーム AI であると予測する理由」として最も多く回答されたのが、「行動の単純さ」であった。本提案手法では、感情的行動モデルの切り替えによって行動傾向の変化が発生している。しかし、ゲームの状況によっては長時間感情決定値が変化しないこともある。表現力不足や報酬設定により、提案手法の各行動モデルは少数の技を多用する傾向がある。これらの要因が重なることで「行動の自然さ」「行動傾向変化の自然さ」の両方で低い評価を受けたと推測する。

SL モデルは本実験では、クラス分類問題として出力された行動を確率 1 でとっていたが、行動の確率分布（いわゆる policy）として確率的な行動を取るなどの工夫はありうるかもしれない。

	SL	RL	SL_N	既存 AI1	既存 AI2	人間
行動の自然さ	Mean	1.33	1.41	1.72	1.58	2.66
	SD	1.12	1.07	1.16	1.41	1.16
行動傾向変化の自然さ	Mean	1.64	1.72	1.35	1.56	2.31
	SD	1.10	1.23	1.16	1.45	1.06

表 4 動画視聴評価実験における各評価値の平均と標準分散

8.2.2 対戦による評価実験

対戦実験について、6 種の評価対象別に平均値と標準偏差を表 5 に示す。

「楽しさ」の項目について、既存 AI の平均が 2.75 に対し、SL モデルは平均 2.33、RL モデルが 3.00 を示している。特に RL モデルは全ゲーム AI の中で高く評価されている。

また、「行動傾向変化の自然さ」について、RL モデルと既存 AI1 の評価値平均について Student の t 検定 ($p < 0.05$) を行った結果、RL モデルの平均値は既存 AI1 モデルの平均値よりも大きく ($p = 0.0079$)、統計的に有意差があることが示された。

「楽しさ」では、RL モデルと既存 AI1 の評価値の平均値

*1 動画撮影に参加した人間は、自身を対戦型格闘ゲーム初級者と認識している男性 1 名である。本評価実験の意図や目的は説明せず、対戦相手を倒す事を目的として対戦して貰った。

では有意差は認められなかったものの、SL_N モデルに対しては、t 検定によって RL モデルのほうが有意に優れることが示された。挙動がはっきりと変化することが、対戦して楽しいと思える要因になっていると判断できる。

「行動傾向変化の自然さ」についても、RL モデルが既存 AI2 モデルに対して有意に優れている。特に、RL モデルの強調された感情的行動による行動傾向の変化を人間が認識することで、ゲームの楽しさに良い影響を与えたと考ええる。また、動画視聴実験では行動傾向変化をやや不自然(1.73)と評価された RL モデルが、対戦実験ではかなり自然(3.30)と評価された点は本研究の大きな驚きである。おそらくなんらかの対戦ならでの知覚変化にうまく適合したものと考えているが、詳細は不明であり、今後分析を行いたい。

		SL	RL	SL_N	既存 AI1	既存 AI2
楽しさ	Mean	2.33	3.00	2.37	2.75	2.70
	SD	1.10	0.86	0.80	0.72	1.30
行動傾向変化の自然さ	Mean	1.45	3.29	1.79	2.45	2.12
	SD	0.86	0.88	0.95	0.86	1.09
行動の自然さ	Mean	1.66	2.12	1.54	2.70	3.04
	SD	0.94	1.13	1.04	0.67	1.09

表 5 対戦評価実験における各評価値の平均と標準分散

9. おわりに

本研究では、対戦型格闘ゲームにおいて、感情を考慮した人間らしいゲーム AI 手法の提案と構築を行った。

まず、被験者にゲーム AI と対戦をしてもらい、状況ごとの行動とそのときの感情を採取した。そして、状況と感情のペア群から、教師あり学習によって感情決定器を構築した。さらに、状況と行動のペア群を、感情によって分割したうえで教師あり学習することで、「喜び」「怒り」「悲しみ/嫌悪」「ニュートラル」の 4 種の感情的行動モデルを作成した。教師あり学習に加えて、より明確な行動傾向の変化を演出するため、特殊な報酬を用いて強化学習による感情行動モデルも作成した。

対戦評価実験の結果、「楽しさ」「行動傾向変化の自然さ」について提案手法が優れていること、強化学習によるモデルのほうがより良い結果となることが確認出来た。ただし、対戦評価実験における「動きの自然さ」、動画視聴評価実験における「動きの自然さ」「行動傾向変化の自然さ」については、両提案手法が既存ゲーム AI に劣る結果を示していた。アンケートの結果からも、行動モデルの人間らしさについては改善の必要があると考える。

今後の展望として、学習手法や報酬設定を吟味することで行動モデルの「動きの自然さ」を向上させることや、本研究では採用出来なかった基本的感情の採用や感情の強度を考慮することで、ゲーム AI の行動傾向により多様さを演出することなどがある。

参考文献

- [1] Sila Temsiriririkkul, et al, Production of Emotion-based Behaviors for a Human-like Computer Player, 17th International Conference on Intelligent Games and Simulation (GAMEON' 2016), pp.49- 54, September 13-15, 2016
- [2] Paul Ekman (1992) An argument for basic emotions, *Cognition and Emotion*, 6:3-4, 169-200
- [3] 木村 知宏, 即時的な反応を要求するゲームとその熟達が感情経験に与える影響, *デジタルゲーム学研究*, 2020, 13 巻, 1 号, p. 21-29
- [4] 飯田弘之, et al. プレーヤー感情のゲーム・パターン依存性. *研究報告エンタテインメントコンピューティング (EC)*, 2011, 2011.3: 1-4.
- [5] LU, Feiyu, et al. Fighting game artificial intelligence competition platform. In: 2013 IEEE 2nd Global Conference on Consumer Electronics (GCCE). IEEE, 2013. p. 320-323.
- [6] <https://github.com/TeamFightingICE/Gym-FightingICE>
- [7] MNIH, Volodymyr, et al. Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602, 2013.
- [8] 織田 弥生, 野 ルリ子, 阿部 恒之, 菊地 賢一, 感情・覚醒チェックリストの作成と信頼性・妥当性の検討, *心理学研究*, 2014, 85 巻, 6 号, p. 579-589