

# 格闘ゲームにおける選択的注意を取り入れたゲーム AI Game AI Adopting Selective Attention in a Fighting Game

柴田 悠真

Yuma Shibata

法政大学情報科学部コンピュータ科学科

E-mail: yuma.shibata.2s@stu.hosei.ac.jp

## Abstract

*Selective attention is the act of paying attention to only a subset of many pieces of information. Players of fighting games quickly defend or counterattack by paying selective attention to a particular action of the opponent. Although non-player characters (NPCs) are important training partners for human players, existing NPCs do not perform selective attention. Therefore, playing against NPCs does not allow players to practice against opponents paying selective attention. We develop AI with selective attention using Q-learning. It is designed to get information quicker than usual if the action of the opponent is the one to which it is paying selective attention and otherwise to get information slower. To combine selective attention with Q-learning, the discount rate of Q-learning and the components of the reward change according to its remaining hit points. We conducted an experiment to examine whether participants could notice the AI's selective attention and identify the attack to which it is paying selective attention. In the experiment, we showed the participants videos of the developed AI's fighting. The results showed that the developed AI had a low ability to counterattack and that it was difficult to judge whether it was paying selective attention to some attack.*

## 1. はじめに

コンピューターゲームには、様々なジャンルのゲームが存在し、格闘ゲームはその中の 1 つである。格闘ゲームは対戦型アクションゲームであり、キャラクターが主に 1 対 1 で相手を攻撃する。キャラクターには体力が設定されており、攻撃によって相手の体力を 0 にした方が、もしくは相手の体力をより減らした方が勝利となる。格闘ゲームではプレイヤーの対戦相手として、ノンプレイヤーキャラクター (NPC) が実装されている。NPC はキャラクターや強さを自由に設定できるため、プレイヤーの練習相手としてよく用いられる。多くのプレイヤーは NPC と練習した後、対人戦に臨むという流れを繰り返す。格闘ゲームで NPC はプレイヤーが強くなって楽しみを感じる上で重要な役割を果たしている。

プレイヤーが NPC との対戦で練習することの 1 つに、対戦時の状況に応じた意識配分の変更がある。格闘ゲームの対戦中は、画面上に多くの情報が存在しているため、全ての情報に対して意識を向けることは難しい。こうした問題に対処するため、多くのプレイヤーは選択的注意を行う。選択的注意とは、多くの情報の中から特定の情報に対して選択的に注意を向けることである。格闘ゲームでは、プレイヤーが相手の特定の挙動に注意を向けることを指す。プレイヤーは選択的注意を行うことで、注意を向けている挙動を相手が取った時、すぐに防御や反撃を行える。

プレイヤー同士の対戦で、相手が選択的注意を向けている挙動を自分が取っても、相手に対処されてしまう。そのため、相手が選択的注意を向けていない挙動を選択することで、相手に攻撃を通すことが重要である。しかし、既存の AI は選択的注意を行わず、全ての挙動に対して同じ反応速度を持ち、自身の攻撃の通りやすさが変化しない。そのため、相手の挙動から自分がした行動に対して選択的注意を向けているかどうか判断できず、特定の攻撃のみが相手に通ることもない。このような選択的注意を実装していない NPC と対戦しても、選択的注意を向けている相手を想定した練習をすることは難しい。こうした練習を実現するには、相手の攻撃に適切に選択的注意を向け、かつその対象がある程度対戦相手から見て分かる AI を開発する必要がある。

本研究では、選択的注意を行うプレイヤーに攻撃を通す練習を NPC 戦でできるようにするために、選択的注意を取り入れたゲーム AI を開発することを目的とする。開発する AI が選択的注意を行っている場合、選択的注意を向ける攻撃を AI に決定させる。選択的注意を向けている攻撃ではない攻撃を相手が取った場合、相手の座標や攻撃を取得するのに時間がかかるようにする。

## 2. 関連研究

John ら [1] は、畳み込みニューラルネットワークとモンテカルロ木探索を組み合わせ、格闘ゲームの AI を開発した。この AI のパフォーマンスはモンテカルロ木探索のみで開発された AI のパフォーマンスには及ばなかったが、格闘ゲーム用の AI を開発に、教師あり学習、強化学習、モンテカルロ木探索の組み合わせが活用できる可能性を示した。Liang ら [2] は、二重深層 Q ネットワーク (DDQN) を用いて、既存の格闘ゲームの AI を開発した。DDQN は、Q 学習で AI が Q 値を過剰に高く見積もる問題を回避するために導入されている。DDQN では、行動決

定のネットワークと Q 値の更新に用いるネットワークを使い分けることで、学習を安定させながら、AI に最適な行動選択をさせられる。これによって開発された AI は、ゲーム内の AI と対戦で 95% 近い勝率を生み出した。

### 3. 準備

#### 3.1. FightingICE

FightingICE [3]は格闘ゲームの AI 開発を目的としたプラットフォームで、立命館大学で開発された。人間の認知能力の限界を模倣するために、対戦相手の情報の取得 15 フレーム (0.25 秒) 以上かかるように設定されている。FightingICE には、あらかじめ 3 つの AI が用意されている。1 つ目は JerryMizunoAI である。これは k-nn アルゴリズムを用いて対戦相手の挙動を予測する AI である。2 つ目は LoadTorchWeightAI である。これはニューラルネットワークを用いて行動を選択する AI である。3 つ目は MctsAi である。これはモンテカルロ木探索を用いて行動を選択する AI である。これら 3 つの AI は、開発した AI が学習する際の対戦相手として用いる。本研究では、これら 3 つの AI を学習用 AI と呼ぶ。また、本研究で用いるのは ZEN というキャラクターに固定し、体力の最大値は 400 とする。

#### 3.2. Q 学習

Q 学習 [4]は強化学習の 1 つで、AI が特定の状態でどの行動を選ぶべきかを学習する手法である。AI は状態  $s$  で行動  $a$  を選択した場合の期待報酬の合計の値を学習する。これを Q 値と呼び、 $Q(s, a)$  と表す。Q 学習を進めると、最終的に AI は全ての状態  $s$  と行動  $a$  の組み合わせに対する Q 値を学習し、Q 値を基に AI が最適な行動を選択するようになる。Q 値は以下の式で更新する。

$$Q(s, a) \leftarrow Q(s, a) + \alpha [R + \gamma \max_{a'} Q(s', a') - Q(s, a)]$$

$R$  は現在の報酬、 $s'$  は次の状態、 $a'$  は次の行動である。 $\alpha$  は学習率といい、この値が大きくなるほど Q 値が更新されやすくなる。 $\gamma$  は割引率といい、この値が大きいほど長期的な報酬を重視し、この値が小さいほど短期的な報酬を重視する。

#### 4. 提案手法

選択的注意を取り入れた FightingICE の AI を開発する。AI の学習手法には Q 学習を基本として用いる。選択的注意を実現するための要素技術として以下の 2 つを提案する。1 つ目は、選択的注意の有無と注意を向けている挙動を Q 学習の状態として入力させるものである。格闘ゲームでプレイヤーは常に選択的注意を行っているわけではないため、AI に対しても選択的注意を行っているかどうかを状態として入力する。AI は選択的注意を向ける挙動を次の手順で決定する。まず、全中距離攻撃の出現回数を表す変数を  $f$  とし、初期値  $f = 0$  として設定する。次に、各中距離攻撃に対してカウント  $C_i$  を設定し、初期値を  $C_i = 0$  とする。対戦中に相手が中距離攻撃  $i$  を行うたびに、 $f \leftarrow f + 1$  に更新し、その値を攻撃  $i$  のカウント  $C_i$  に加える。すなわち、攻撃  $i$  のカウントの更新は以下の式で定義される。

$$C_i \leftarrow C_i + f$$

その時点におけるカウント  $C_i$  の値が最も大きい攻撃に対して選択的注意を向ける。これによって、相手の攻撃パターンが変化した時も、素早く選択的注意を向ける攻撃を切り替え、その攻撃に対して対応できるようになる。

Q 学習における状態は、1 つの値として管理する。この値は、相手の座標、相手の挙動、相手に与えたダメージと自身が受けたダメージのどちらであるか、選択的注意の有無の 4 つの状態を数値化したデータから成り立っている。相手の  $x$  座標と自身の  $x$  座標の差の絶対値を  $x$ 、相手の  $y$  座標から自身の  $y$  座標を引いた値を  $y$ 、相手の挙動を数値化した値を  $a$ 、相手に与えたダメージと自身が受けたダメージのどちらであるかを数値化した値を  $b$ 、選択的注意の有無を数値化した値を  $c$  とすると、状態の値  $s$  は以下の式で表される。 $b$  は、相手に与えたダメージなら 0、自身が受けたダメージなら 1 の値を取る。 $c$  は、選択的注意を行っているときは 0、行っていないときは 1 の値を取る。

$$s = 10000000x + 10000y + 100a + 10b + c$$

状態の値  $s$  に対して、一意に挙動の組が定まっている。これを Q 学習における行動とする。挙動の組は、空中で可能な挙動の組と地上で可能な挙動の組のどちらかに分かれている。自身が空中にいるかどうかで挙動の組が異なる。自身の行動を決定する際は、現在の状態で可能な行動のうち、最も Q 値が高い行動を選択するようにしている。

提案する 2 つ目の要素技術は、選択的注意の有無に応じて、情報の入力に遅延をかけるものである。選択的注意を向けている挙動を相手がとった場合、15 フレームで対戦相手の情報を取得する。選択的注意を向けていない挙動を相手が選択した場合、21 フレーム (0.35 秒) で対戦相手の情報を取得する。選択的注意を行っていない場合は、18 フレーム (0.30 秒) で対戦相手の情報を取得する。

選択的注意と Q 学習を両立させるため、自身の残り体力に応じて割引率と報酬の構成要素の比率を変更する。選択的注意は相手の次の攻撃に対処するために行うものであるため、将来の価値よりも現在の価値を重視する。そのため、選択的注意を行っているときは割引率を低くする。しかし、格闘ゲームでは対戦相手に勝利するという長期的な価値を重視する必要があるため、割引率が低いままだと Q 学習に悪影響が出る可能性がある。そこで、Q 学習と選択的注意を両立させるために、体力が少なくなるほど割引率が低くなるようにする。こうすることで、自身の体力に余裕があるときは Q 学習を行い、余裕がないときは選択的注意を行うようにする。割引率は、戦闘開始時の体力を 1 とした時の現在の体力の割合となる小数にする。報酬の構成要素の比率も体力に応じて変更する。格闘ゲームでは、自身が有利な状況では攻勢に転じ、自身が不利な状況では守勢に転じることが重要であるとされる。そのため、自身の体力に対して相手の体力が少ないときは、自身の体力が減ることより相手の体力を減らすように報酬を設定し、自身の体力に対して相手の体力が多いときは、相手の体力を減らすことより自身の体

力を減らさないように設定する。また、AI が防御ばかりして攻撃しなくなることを防ぐため、自身の数値に 1 を加えて、互いにダメージを受けない挙動は小さなペナルティになるようにする。互いの現在の体力の合計を 1 とした時の自身の現在の体力の割合にあたる小数を  $a$ 、相手の現在の体力の割合にあたる小数を  $b$ 、自身の行動によって減少した相手の体力を  $c$ 、減少した自身の体力を  $d$  とする。この時、報酬  $r$  は以下の式で表される。

$$r = (a \times c) - (b \times d + 1)$$

## 5. 実装

Q 学習の状態を用いる値に、相手の  $x$  座標と自身の  $x$  座標の差の絶対値  $x$  と、相手の  $y$  座標から自身の  $y$  座標を引いた値  $y$  を用いるが、この値をそのまま用いると状態数が膨大になり、膨大な計算量が必要となる。状態数を抑えるため、相手と自身の  $x$  方向の距離を短距離と中距離と長距離の 3 つに分類し、それぞれの距離で  $x$  と  $y$  の値を調節する。短距離では、 $x$  と  $y$  の値を 2 の倍数に丸める。中距離では、 $x$  と  $y$  の値を 3 の倍数に丸める。長距離では、 $x$  の値を中距離と長距離の境目となる値とし、 $y$  の値を 0 とする。

## 6. 実験

### 6.1. 目的

開発した AI が選択的注意を適切に行えているかを客観的に判断するためにアンケートを行う。アンケートの被験者には格闘ゲームの上級者と中級者を選定する。アンケートでは、被験者に 10 個の対戦動画を見せ、AI が選択的注意を向けているかどうかと、どの攻撃に対して選択的注意を向けているかを被験者に尋ねる。上級者と中級者のそれぞれの回答での正答率を元に、選択的注意がどのくらい適切に行えているかを判断する。

### 6.2. アンケートの選択肢

アンケートでは、2 種類の選択肢を用意した。本研究では、選択的注意を向けている攻撃を選択する際に用いる選択肢を攻撃の選択肢と呼び、選択的注意を始めた時または切り替えた時の自身の残り体力を選択する際に用いる選択肢を体力の選択肢と呼ぶ。

攻撃の選択肢は、3 つの攻撃に「分からない」を加えた 4 択である。3 つの攻撃は、FightingICE で ZEN が使える攻撃の中で、中射程以上の距離を持つ攻撃である。1 つ目の攻撃は、エネルギーを少量消費し、低姿勢で足を伸ばしながら移動する攻撃であり、スライディングと呼ぶ (図 1(a))。2 つ目の攻撃は、前方に低軌道のジャンプをしながら殴りかかる攻撃であり、飛び込み攻撃と呼ぶ (図 1(b))。3 つ目の攻撃は、エネルギーを大量消費し、キャラクターの腰の位置あたりの高度を維持し続けながら前方に長距離移動する弾を放つ攻撃であり、波動拳と呼ぶ (図 1(c))。

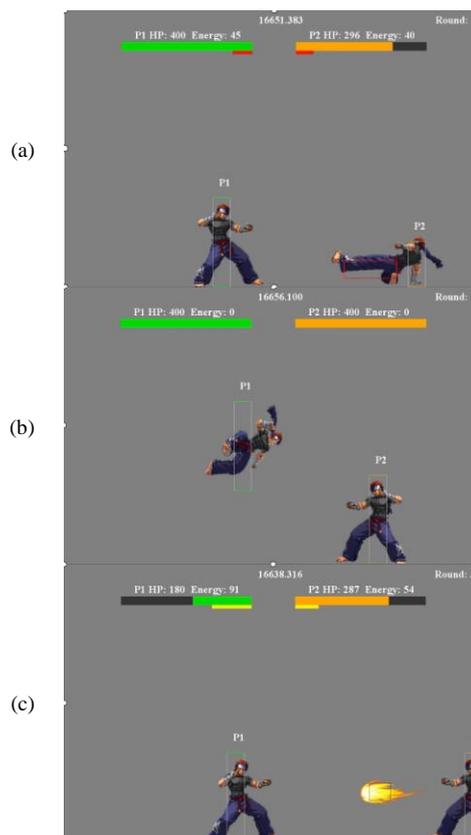


図 1 (a)スライディング, (b)飛び込み攻撃, (c)波動拳のモーション

体力の選択肢は、自身の残り体力が 75%、50%、25% の 3 択に「分からない」を加えた 4 択である。本実験では戦闘開始時の体力が 400 であるため、残り体力が 75% の時の数値は 300、50% の時の数値は 200、25% の時の数値は 100 である。体力の選択肢では、体力の割合と数値が併記されている。

### 6.3. 対戦動画での相手

学習用 AI は、対戦動画での AI の対戦相手としては不適なため、動画用に AI を開発した。これを本研究では動画用 AI と呼ぶ。学習用 AI が動画に不適な理由は 3 つある。1 つ目は、攻撃のパターンが変化しない点である。開発した AI は相手の攻撃パターンの変化に対応する。しかし、学習用 AI は攻撃のパターンが変化しないため、対戦の動画を見ても AI が攻撃パターンの変化に対応できているかの判断ができない。2 つ目は、エネルギーを節約しない点である。エネルギーを消費する波動拳やスライディングを連発するためには、なるべくその他の技でエネルギーを消費しないような立ち回りが必要である。しかし、学習用 AI は対戦で勝利するためにエネルギーを随時消費するため、波動拳やスライディングを使う回数が減り、選択的注意を向けているかの判断が難しくなる。3 つ目は、相手から近距離の位置を維持する点である。近距離では選択的注意を向けていても反応が間に合わないため、動画では中距離からの攻撃を多めにすることが

重要である。しかし、FightingICE では中距離技の数よりも近距離技の数の方が遥かに多いため、学習用 AI は中距離から攻撃するよりも相手に近づくことを優先する。

動画用 AI は特定の攻撃を中距離で優先して使い、対戦相手の体力が一定以下になったらその攻撃を切り替えることができる。また、自身が選択可能な攻撃からエネルギーを消費する攻撃を極力減らしているため、エネルギーを節約することもできる。また、自身が近距離にいる際は一定確率で距離を取ることで、中距離攻撃の回数も増え、選択的注意を向けているどうか判断しやすくなる。

#### 6.4. アンケートの種類

アンケートは①～④の 4 つに分かれている。アンケート①では、対戦動画を見て、AI が選択的注意を向けている攻撃を選択肢から答える。動画は 2 つあり、動画の最初から選択的注意を向けている。また、選択的注意を向けている攻撃は、対戦動画の中では変化しない。アンケート②では、対戦動画を見て、AI が選択的注意を始めたタイミングを体力の選択肢から答えて、選択的注意を向けている攻撃を選択肢から答える。動画は 2 つあり、AI の体力が一定値を下回ると選択的注意を開始する。アンケート③では、対戦動画を見て、AI が選択的注意を切り替えたタイミングを体力の選択肢から答えて、切り替える前と後で選択的注意を向けている攻撃を選択肢から答える。動画は 6 つある。アンケート④では、「AI が相手の攻撃にどれくらい対処できていたか」「相手の攻撃にさらに対処できるようにすればいいか」「気づいた点、気になった点」について自由記述で答える。

#### 6.5. 手順

被験者として格闘ゲームの上級者 3 人と中級者 4 人を募集した。実験ではまず、アンケートの流れを説明する。次に、アンケートに答える上で必要な情報について、FightingICE のゲームシステムと選択的注意を説明する。次に、選択的注意を向ける攻撃のモーションを被験者に見せて、その主な対処法を説明する。最後に、開発した AI の特徴について説明する。これらの後、FightingICE の通常の対戦を 3 試合見てもらってから、被験者にアンケートに回答してもらった。アンケート①～③について、説明後、被験者にアンケートに回答してもらおうという流れを繰り返した。最後にアンケート④を行った。

#### 6.6. 結果

アンケート①での上級者と中級者の問題別の正答率は表 1 の通りである。アンケート②での上級者と中級者の問題別の正答率は表 2 の通りである。体力とは、選択的注意を始めた時の自身の体力のことであり、攻撃は選択的注意を向けていた攻撃のことである。アンケート③での上級者と中級者の問題別の正答率は表 3 の通りである。攻撃前とは、始めに選択的注意を向けていた攻撃のことであり、体力とは選択的注意を切り替えた時の自身の体力のことであり、攻撃後とは選択的注意を切り替えた後に、選択的注意を向けていた攻撃のことである。攻撃の選択肢ごとでの正答率と、体力の選択肢での正答率をまとめると表 4 の通りである。

表 1 アンケート①の正答率

	上級者	中級者
動画 1	66.7%	25.0%
動画 2	0.0%	25.0%

表 2 アンケート②の正答率

	上級者		中級者	
	体力	攻撃	体力	攻撃
動画 3	33.3%	0.0%	25.0%	25.0%
動画 4	66.7%	0.0%	50.0%	50.0%

表 3 アンケート③の正答率

	上級者			中級者		
	攻撃前	体力	攻撃後	攻撃前	体力	攻撃後
動画 5	66.7%	66.7%	0.0%	25.0%	25.0%	25.0%
動画 6	33.3%	0.0%	66.7%	100.0%	0.0%	75.0%
動画 7	33.3%	33.3%	0.0%	25.0%	0.0%	25.0%
動画 8	100.0%	0.0%	66.7%	100.0%	0.0%	50.0%
動画 9	0.0%	100.0%	0.0%	100.0%	25.0%	100.0%
動画 10	0.0%	0.0%	66.7%	75.0%	0.0%	50.0%

表 4 各攻撃と体力の正答率

	上級者	中級者
スライディング	33.3%	45.8%
飛び込み攻撃	20.0%	60.0%
波動拳	53.3%	60.0%
体力	33.3%	18.8%

### 7. 議論

アンケート①は質問項目が 1 つしかないため、被験者が注目すべき箇所も少ない。しかし、その正答率は想定よりも低かった。この理由として、被験者の目が慣れていなかったことが考えられる。実際にアンケート④では、「実験の後半になるにつれて目が慣れてきた」といった意見が多かった。本実験では、実験が長引いて被験者の集中力が落ちないように、アンケート前に見せる対戦動画の数を 3 個に絞ったため、アンケート①の時点では目が慣れていなかった可能性が高い。

アンケート②では、上級者は中級者に比べて、選択的注意を始めたタイミングの正答率は高かったが、選択的注意を向けている攻撃の正答率は 0% だった。一方で、中級者の正答率は両方の項目で一致していた。このようになった理由として、アンケート②の段階では、上級者は自分たちの基準で選択的注意を向けている攻撃を判断していた可能性が考えられる。開発した AI は学習用 AI よりも弱いくらいの性能であるため、上級者が想定する立ち回りとはかけ離れた立ち回りを取っていた可能性が高い。

アンケート④で上級者の 1 人が「アンケート③の序盤になってようやく開発した AI 基準で考えることができた」と回答していた。実際に、動画③の攻撃の選択肢での上級者 3 人の回答は一致しており、動画④でも 3 人中 2 人の意見は一致していた。アンケート③では、上級者は選択的注意を向けている攻撃の正答率が低く、中級者は選択的注意を切り替えたタイミングの正答率が低かった。

前者の理由としては、開発した AI は反撃せずに回避することが多く、上級者目線では選択的注意を向けている攻撃に対処できているかの判断が難しかったことが考えられる。後者の理由としては、2 回の動画視聴では、中級者は選択的注意を切り替えるタイミングまで絞り込めないことが考えられる。アンケート④では中級者の 2 人が「選択的注意の対象が変わっているのは分かったが、いつの間にか変わっていた」と回答している。

表 4 からは上級者は飛び込み攻撃の正答率が低く、中級者はスライディングの正答率が低かった。前者の理由としては、飛び込み攻撃は反撃が簡単な攻撃であるのにもかかわらず、上級者にとっては反撃が少ないと感じたことが考えられる。後者の理由としては、スライディングは発生が早く、中級者にとっては対処できていないと感じる場面が多かったことが考えられる。

## 8. おわりに

本研究では、Q 学習を用いて格闘ゲームでの選択的注意を取り入れたゲーム AI を開発した。しかし、格闘ゲームの上級者と中級者でも、AI の選択的注意の対象を見極めるのは難しかった。

今後の課題として、AI の反撃性能を高めることが挙げられる。このためには、Q 学習の報酬の見直しが必要である。現時点では、AI は行動可能なタイミングで必ず行動するようになっているが、反撃する際には攻撃するタイミングまで待機することが必要である。そのため、AI に反撃させるには、選択的注意の開始から反撃までの一連の流れを報酬として設定しなおす必要があると考える。また、今回の実験では、被験者が FightingICE に慣れていないことが実験結果に影響した可能性がある。より正確な実験結果を得るには、被験者に FightingICE に事前に触れてもらった上で実験を行うことが重要である。

## 文 献

- [1] J. P. Q. Tomas, A. N. De villa, N. j. R. Aguas and J. R. G. Lim, "Developing an Adaptive AI Agent using Supervised and Reinforcement Learning with Monte Carlo Tree Search in FightingICE," *Proc. CIIS*, pp. 31-36, 2021.
- [2] L. Hai and L. Jiaqi, "A Study on the Agent in Fighting Games Based on Deep Reinforcement Learning," *Mobile Information Systems*, vol. 2022, no. 9984617, pp. 1-8, 2022.
- [3] R. U. Intelligent computer entertainmentlab., "FightGame AI Competition," [Online]. Available: <https://www.ice.ci.ritsumeai.ac.jp/~ftgaic/>.
- [4] C. J. C. H. Watkins and P. Dayan, "Q-Learning," *Machine Learning*, vol. 8, pp. 279-292, 1992.