

## 機械学習を用いたゲーム AI の人間らしい行動に関する研究

三浦 大輔<sup>†</sup> 會澤 邦夫<sup>†</sup>島根大学大学院総合理工学研究科<sup>†</sup>

## 1. はじめに

近年、コンピュータゲームにおいて「e スポーツ (e-sports)」が人気となってきている。e スポーツは、コンピュータゲームを通して人と人が競い合う対人戦を目的とした、競技性の高いゲームとして確立されてきた[1]。しかし、初心者が家庭でゲームをプレイする上では、操作を覚えるまでは対人戦ではなく、コンピュータが操作するキャラクター (COM) と対戦することが多い。そのため、プレイヤーが e スポーツのようなゲームの本来の醍醐味である対人戦に興味を持つためには、COM との対戦においても面白いと感じる現実的な強さのゲーム AI が求められる。このゲーム AI が理不尽な動作によりプレイヤーにとって不快感を覚えるものであれば、プレイヤーがゲームから離れてしまうという問題が発生してしまう。そこで、本研究では、COM と対戦する人にも楽しむことができる、対人戦をしていると感じられるような「人間らしい」ゲーム AI が実現できれば、プレイ人口も増えてより活発になるのではないかと考える。

## 2. 人間らしいゲーム AI

これまでも多くのゲーム AI の研究はされており、コンピュータの処理速度の向上や、技術の発展により、囲碁や将棋などでも人間に勝利するまでにゲーム AI は発展してきている。しかし、それらは強さを追求したゲーム AI であるために、行動は最適化されており、人間にとっては機械的で人間らしくないと感じてしまう。e スポーツなどの対戦型アクションゲームにおいて最適化された振る舞いは、人間の反応速度を遥かに上回るなど、ゲームとしてプレイヤーの興味をなくす原因となってしまふ。COM の人間らしい振る舞いの研究においては、人間の行動規則を用いた手法が研究されている[5]。本研究ではこれらの手法に改良を加え、より人間らしいと感じられるゲーム AI を追求することが目標である。

## 3. 人間らしさとは

これまでも人間らしさの研究はされてきたが、開発者の主観が関わってしまう。したがって、人間らしさの明確な定義はないため、本研究における人間らしさとは、ゲームをプレイしているプレイヤーにとって「ゲーム AI が人間らしい振る舞いをしているようにみえる」と定義する。

## 4. 実験環境

本研究では e スポーツとしても人気の、対戦型アクションゲームの一ジャンルである三人称アクションシューティングゲーム (TPS) を想定して扱う。TPS では、複数人がゲーム内のキャラクターを操作して対戦を行う。本研究では、TPS の持つ移動に着目した、キャラクターがフィールド内を移動する追跡ゲームを参考に作成したものを扱うこととする[4]。

## 5. 実験用ゲームの概要

本研究で用いるゲームには、アクションゲームを簡易モデル化した追跡ゲームを扱う。このゲームでは、壁で囲まれた、縦 1200 ピクセル、横 1000 ピクセルのフィールド内で逃亡側一体と追跡側二体のキャラクターが存在する。ゲームの終了は、逃亡側が追跡側に捕まるか、制限時間を経過した場合である。プレイヤーはキャラクターの移動にゲームパッドを用いて上下左右と斜めの八方向で操作する。ゲームの実行画面を図 1 に示した。

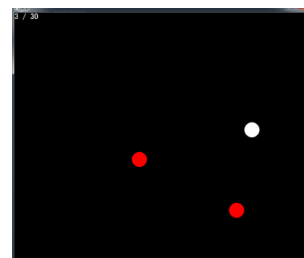


図 1 ゲーム画面

## 6. 実験用ゲーム AI

実験で用いるゲーム AI には強化学習と呼ばれる手法を用いる。強化学習を用いることについては、コンピュータゲームでは大量の教

師データを用意することが困難であるためである。そして、強化学習の手法の中でも、状況を報酬によって設定できる、Q 学習を用いる [2]。

## 7. Q 学習

Q 学習は、状態 $S_t$ のときに、行動 $a_t$ をとった時の行動価値である Q 値、 $Q(S_t, a_t)$ を持つ。Q 学習では、この Q 値を基に行動を選択する。もっとも簡単な一ステップの Q 学習は、次のように定義される。

$$Q(S_t, a_t) \leftarrow Q(S_t, a_t) + \alpha \left( r_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, a_t) \right) \quad (1)$$

現在の状態 $S_t$ で、行動 $a_t$ を選択肢し、次の状態 $S_{t+1}$ に遷移したとき、報酬 $r_{t+1}$ が与えられ、Q 値が更新される。学習率 $\alpha$  ( $0 < \alpha < 1$ ) 割引率 $\gamma$  ( $0 < \gamma < 1$ ) はパラメータである。

## 8. 意識や自我

追跡ゲームに使用するゲーム AI は、プレイヤーが一人の横スクロールアクションゲームに Q 学習を用いて、「ゆらぎ」、「遅れ」、「疲れ」などの人間の行動原則を報酬として用いたものをベースに作成した [5]。この手法によって、人間が操作しているかのようなヒューステリックな振る舞いが可能となる。本研究では、プレイヤーが複数存在する対戦型のゲームである追跡ゲームを用いるため、相手の人間らしいヒューステリックな動作の観測を、自分のゲームプレイにおいてゲームプレイ中に直感的に感じ取ることができる。そのため、人間らしくみえるのに欠かせない要素としての意識や自我といった要素を考えることとした。意識や自我に関して、具体的な研究というのはまだあまり進んでいないが、「言語」、「感情」、「倫理」、「ユーモア」、「常識」の五つの要素は、それらに欠かせない要素であると考えられている。「言語」「感情」は人間のコミュニケーション手段、「倫理」は人に害をなすのを防ぎ、「ユーモア」は人を楽しませる、「常識」は個人レベルを超えたグループなどのレベルで共有されている知識のことをいう。本研究である追跡ゲームにおいては、ゲーム上ということで、関連性の低い、言語、感情、倫理を除いたユーモアと常識に重点をおき、追跡ゲームのルールや、人がプレイした結果どう振る舞うのかという点を常識として設定してやり、常識にない振る舞いをユーモアとしてゲーム AI に理解させることで、COM が人間らしさに必要な意識や自我を持っているようにみえるゲーム AI の実現を目指した。

## 9. 考察とまとめ

本研究では、人間の行動原則の制約を課す手法を用いた Q 学習での追跡ゲーム AI に、人間らしさを意識や自我の存在をもつ人間らしいゲーム AI 目指した。追跡ゲームは、移動のみを扱ったものであったために、主観的な評価としての人間らしく見えるゲーム AI に、その要素である意識や自我をもたせることで、より人間らしくとされるゲーム AI を作成した。今後の展望としては、移動以外の要素が存在するゲームにおいて言語や感情の要素を持つことが可能になると考えられるので、その要素についても検証したい。

## 参考文献

- [1] INTERNATIONAL E-SPORTS FEDERATION, <http://ie-sf.org/eng/index.asp>. (2016年9月29日)
- [2] Watkins, C Learning from Delayed Rewards, PhDthesis, Cambridge University, Cambridge, England, 1989.
- [3] Warrens S. McCulloch, Walter Pitts, A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, Vol. 5, No. 4, pp. 115-133, 1943.
- [4] 安武 諒, 山口 崇志, マッキン ケネス ジェームス, 永井 保夫, チューリングテストによるゲーム AI の客観的評価, 東京情報大学研究論集, Vol. 16 No. 1, pp.33-46, 2012.
- [5] 藤井 叙人, 佐藤 祐一, 若間 弘典, 風井 浩志, 片寄 晴弘, 人間の行動原則の制約下で自動獲得されたビデオゲーム COM プレイヤーの「人間らしい」振る舞いの主観評価, EC2013, pp.26-33, 2013.