

Believable Agent の生体情報による評価

森田 和貴^{1,a)} 橋山 智訓^{1,b)} 田野 俊一^{1,c)}

概要: 近年, コンピュータゲームの分野において人間らしいと判断されるような Believable Agent の研究が盛んに行われている. Agent の Believability が向上すると, プレイヤの没入感が高まり, ゲームを高品質にすると考えられている. Believability は人間の評価に依存する主観的な概念であり, 評価することは難しい. Believable だと判断される Agent が, 実際にどのように解釈されているかは未だに不明である. Agent を構築するための明確な評価指標は存在しない. 本研究では, Believable Agent を Ad hoc に構築し, その Agent を評価する人間, 操作する人間の生体情報を観察することでその変化と Believability の関係について検討する.

Evaluation of Believable Agent by biological information

KAZUKI MORITA^{1,a)} TOMONORI HASHIYAMA^{1,b)} SHUNICHI TANO^{1,c)}

Abstract: In recent years, computer games have been developed greatly both in commercial and research field. The focus of the research is widely spread not only making strong AI but also giving player a lot of fun, and so on. Believable Agent is an agent that makes players feel human-like or does not make players feel suspicious. Improvement of the Believability of the agent increases the immersive feeling to the game. As a result, the computer games will have higher quality. However, Believability is quite subjective and depends on each human evaluation. It is hard for human beings to evaluate the believability of the agent. Evaluating subjective concepts with quantitative indicators includes the players activities such as key input frequency, acquiring the state in the game, and biological information. In the systems using Biological information, the biological signals such as electroencephalogram, skin conductance activity, gaze and pulse, are transformed into meaningful information for the players. Studies of computer games using biological information are mainly used for the purpose of improving the experience and pleasure. There are few studies which focus on Believability. In this research, we try to use biological information technology to evaluate agents Believability.

1. はじめに

近年, 人の代わりを果たす AI の必要性が増している. コンピュータゲームの分野では, AI の発展が目覚ましく, 囲碁や将棋, ポーカーなど様々なゲームで AI が人間に勝利する事例が報告されている. その一方で AI があまりにも人間らしくない挙動や思考の仕方をするにより, ゲームの進行に不自然さを感じる報告も増えている. このよう

な現象に対して, 人間と機械をわからないように配置し, それぞれを比較することでどちらが人間であるように感じたかを測る Turing Test が古くから行われている. コンピュータゲームにおいては, Believability という概念として自然で人間らしい挙動を見つける研究が盛んに行われている [1]. AI の Believability が向上すると, 没入感が高まり, ゲームをより高品質にすると考えられている.

Believable Agent を構築するアプローチには, 生物学や心理学に則ったモデリング [2][3] や, ルールベースによる実装などがある [4]. これらのアプローチは, Believability を向上させる一面を見つけることに成功しているが, 未だに Believability を解明するには至っていない. Believable Agent を構築するには, どのような行動が Believability を

¹ 電気通信大学大学院情報理工学研究所
Graduate School of Informatics and Engineering, The University of Electro-Communications

a) k_morita@media.is.uec.ac.jp

b) Hashiyama.Tomonori@uec.ac.jp

c) tano@is.uec.ac.jp

向上させており、その行動を取る際にどんな反応が起きているかを解釈することが求められることが考えられる。

Believability を客観的指標を用いて解釈する手法には、入力したキーの頻度や連続的に入力される割合や、ゲームプレイ時の生体情報を測る手法がある。生体情報を取得するアプローチは、脳波や皮膚電気活動、視線や脈拍などの生体情報を計測したいくつかの事例がある [5][6][7]。ただし、これまでの研究では、主に体験や楽しさを向上させる目的で使用されており、Believability に注目した研究は少ない。そこで本研究では、AI と生体情報に着目し、プレイ時の生体情報に加え Believability が向上する振る舞いを行う際の生体情報の変化を観察する。

対象は、横スクロール型 2D アクションゲームで Believable Agent を構築する目的の大会が行われた実績を持つ、Infinite Mario Bros の亜種である Infinite Tux を用いる。実際のゲーム画面を図 1 に示す。



図 1 Infinite Tux のゲーム画面

2. 関連研究

2.1 Believable Agent

これまで、人間だと判断されるような, Believable Agent を構築する目的の様々な研究が行われてきた。藤井らは、人間が物事を認識する際に生じる「ゆらぎ」や、認識から行動する際の「遅れ」、行動を続けることによる「疲れ」などの、人間が生物学的に持っている制約を導入することで、人間らしい AI を構築することを試みた [2]。このような AI を構築し、人間らしさを競う大会が何度も行われている。AI を評価する手法には、Turing Test が行われている。この評価方法は、AI が大まかに人間らしいと判断されていることを確かめることができるが、どのような理由で行動を取ったのかという理由が不明なため、何故人間らしいのかという疑問を解決できない。Camilleri らは、ゲームのキャラクタから生じる人間らしさから、ゲームのコンテンツであるレベルに焦点を当てるべきだと主張し、解析を行っている。今回は、ゲームのコンテンツから、ゲームを操作する人、ゲームを観察する人に注目を広げ、観察を行う。

2.2 生体情報

生体情報を用いたゲームに関する研究は、エンタテインメント性に注目したいくつかのアプローチがある。Ambinder

らは、表情や視線、心拍や皮膚電気活動などを測定し、ゲームデザインに利用することで従来の感情状態を推定しながら難易度を変更するシステムよりも満足度が高まった結果を示した [5]。計測した生体情報をゲームの演出に利用する研究の成果は少なくない [6][7]。代蔵らは、動画鑑賞のような受動的なエンタテインメントに対しても、生体情報を用いた演出が有効であることを示している [8]。これらの結果から、Believability を向上させる特徴的な行動を観察すると、生体情報に何らかの顕著な変化が起こることを期待する。

3. 予備実験

3.1 実験条件

予備実験では、人間らしいと判断される行動時の生体情報を計測するために、人間の特徴的な行動を取り入れた Agent とそれ以外の Agent に対しての比較を依頼した。従来研究の評価方法 [9][10] に従い、2つの動画の対を表示し、人間が操作しているように感じた動画を選択する、Turing Test 形式の実験を行った。実験の様子を図 2 に示す。

Q1. 左側 (A) と右側 (B) のどちらが人間が操作している可能性が高いと思いますか？

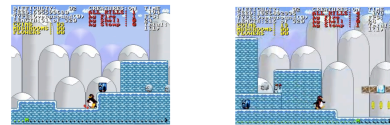


図 2 予備実験の様子

被験者は、20 歳から 26 歳の 12 名であった。1つの動画の時間は、簡単なレベルを達成するのに要する時間とされる 30 秒を上限とした。被験者は、より人間が操作しているように感じた動画を選択するか、どちらも人間であるように感じたか、どちらも人間ではないように感じたかという 4つの選択肢から回答を行った。比較する Agent は、殆ど横スクロールアクションゲームを行ったことがない初心者相当の人間、普段からよく遊ぶ上級者相当の人間の 2名の操作の様子と以下の 3つの NPC の動作ビデオを用いる。1つ目の Agent は、Baumgarten らの A*アルゴリズムを用いた Agent (A*) [11] である。この Agent は、2009 年に実施された Mario AI Competition で優勝を果たした、非常に高度な操作を行う Agent である。2つ目の Agent は、常に前方に進み、障害物を検知した際にジャンプアクションを起こす Forward Agent (FA) を用いる。この Agent は、非常にシンプルな Agent として、しばしば比較対象として用いられる。3つ目の Agent は、Forward Agent をもとにした、条件を満たした際に特徴的な行動を起こすルールベースの Agent (RB) を用いて構築した。特徴的な行動

は、人間の操作を観察した中で、特に目立った行動として以下の5つを用いる。1つ目のルールと行動は、「アイテムブロックを確認した場合、それを全て破壊する」という行動である。2つ目のルールと行動は、「アイテムを確認した際は全て取りに向かう」という行動を取り入れた。アイテムを取りに向かう行動は、2010年に実施された Mario AI Championship の Turing Test Track において優勝を果たした Agent である、REALM を制御する4つの行動の1つであるパワーアップアクションとしても取り入れられており、この特徴点が人間らしいと評価されている可能性がある。3つ目から5つ目までの行動は、「段差を一段ずつ超える」「敵を確認した際は前進を止める」「障害物を1つずつこなす」という初心者の行動によく現れた行動を取り入れた。このように、人間のプレイヤーの操作動画（初心者、上級者）2つと NPC3つ（RB, A*, FA）の合計5つの Agent を、重複しない対10組を作り、動画の左右、対の順番、動画の内容を10種類の中からランダムで選択した。

3.2 実験結果

結果は、RB が、FA, A*, 上級者より人間らしいと判断された。総合的な評価の結果を図3に示す。

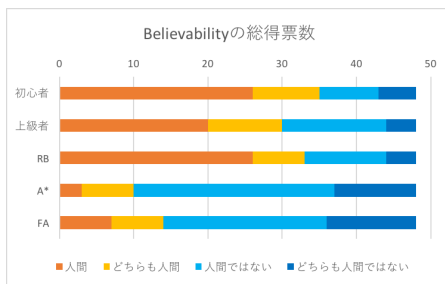


図3 予備実験の結果

実験結果に対してマンホイットニーのU検定を行った結果を表1に示す。順位は[9]に基づき、明確に人間だと判断された場合1、されなかった場合0を適用した。 $p < 0.01$ の範囲で有意差がみられた対を太字で表す。初心者、上級者、RBのAgentがA*に対して有意に人間らしいと判断され、上級者のみがFAに対して有意に人間らしいと判断されなかった。

表1 各対にマンホイットニーのU検定を行った結果

Agent	初心者	上級者	RB	A*	FA
初心者	-	0.0707	0.3519	0.0067	0.0010
上級者	-	-	0.4840	0.0025	0.0515
RB	-	-	-	0.0006	0.0004
A*	-	-	-	-	0.0721
FA	-	-	-	-	-

3.3 考察

実験の結果から、アイテムを取りに向かうような行動をはじめとするいくつかの特徴を取り入れたAgentが人間らしいと判断された。今回のAgentは、人間らしくないと判断されたFAをベースとしているため、アイテムを見つけていない時など大半の時間は単純な行動しか起こさない。それに関わらず、人間の上級者よりも人間らしいと判断されてしまった。被験者に対するアンケートからも、上手い操作が人間らしくないと述べられており、人間が人間らしくないと判断される Turing Test の危うさを示す結果となったと考えられる。

4. 生体情報の観察

予備実験をもとに、ゲームを操作する人の生体情報を観察した。取得した生体情報として、キー入力の圧力、ガルヴァニック皮膚反応（GSR）の2つを取得した。

4.1 ガルヴァニック皮膚反応

ガルヴァニック皮膚反応は、皮膚の電気抵抗値として取得され、緊張状態として計測される。初心者と上級者のゲームプレイ中のGSR値を図4、図5に示す。thresholdは、ゲーム開始前に500回測定した平均値である。

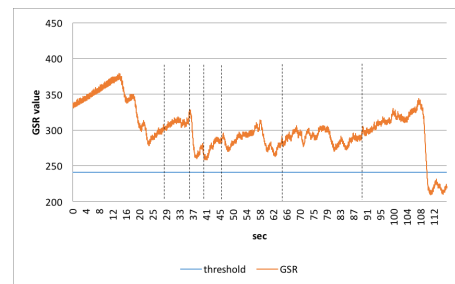


図4 初心者のゲームプレイ時のGSR値

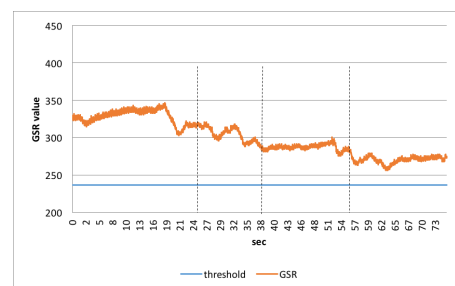


図5 上級者のゲームプレイ時のGSR値

図4から、初心者のGSR値は激しく変化し、上級者は緩やかな変化が起きたことがわかる。点線で示される位置は、今回の予備実験で用いた特徴的な行動を行った瞬間を表す。特徴的な行動の1, 2であるアイテムを取得しに向かう行動を行った際はGSR値が上昇し、特徴的な行動4

のような敵と対面する際は GSR 値が下降していることがわかる。一方、上級者は特徴的な行動を行った際にも変化が見られなかった。このことから、GSR 値が変動している際の行動が人間らしさを向上させている可能性がある。

4.2 圧力

今回の題材とした Infinite Tux は、左右、下のキーとジャンプボタン、ダッシュ/ファイアボタンの5つの入力で制御される。今回は、ジャンプボタンとダッシュ/ファイアボタンの入力に対する圧力を取得した。初心者と上級者の圧力を図 6、図 7 に示す。

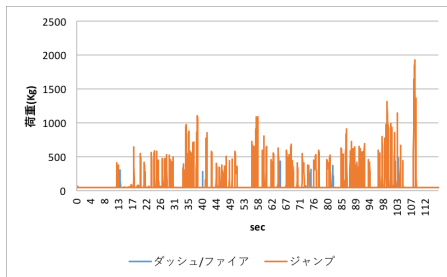


図 6 初心者のゲームプレイ時のボタンの圧力値

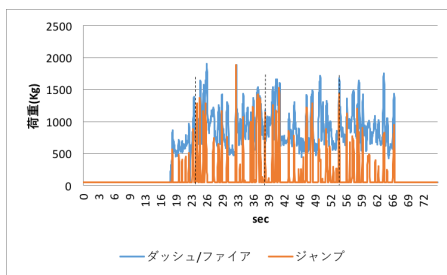


図 7 上級者のゲームプレイ時のボタンの圧力値

今回、予備実験で用いた特徴的な行動時の圧力に、特徴的な行動時以外の圧力との変化は確認できなかった。ただし、初心者はダッシュ/ファイアボタンをほぼ入力しておらず、上級者はほぼ全ての時間で入力し続けていることがわかる。

5. まとめと今後の課題

近年、AI が日常生活の中で必要となる場面が増えており、AI の精度が向上する一方で不自然さを感じる場面が増えてきている。Believability に関する研究分野では、人間らしさを感じるような振る舞いを行う Agent の開発が行われているが、何故人間らしいと判断されているかは解釈されていない。解釈を行う一歩として、人間が行う特徴的な行動を取り入れた Agent を実装し、その特徴的な行動時を観察する際の生体情報の測定も行った。今後は測定する種類を増やし、測定をより細かく分析し、測定した結果と人間の特徴的な行動との関係をより深く検討する。

参考文献

- [1] Yannakakis, Georgios N., and Julian Togelius. "A panorama of artificial and computational intelligence in games." IEEE Transactions on Computational Intelligence and Games 7.4 (2015): 317-335.
- [2] 藤井叙人, et al. "生物学的制約の導入による「人間らしい」振る舞いを伴うゲーム AI の自律的獲得." ゲームプログラミングワークショップ 2013 論文集, pp. 73-80, 2013
- [3] Kersjes, Hanneke, and Pieter Spronck. "Modeling believable game characters." Computational Intelligence and Games (CIG), 2016 IEEE Conference on. IEEE, pp. 1-8, 2016.
- [4] Bojarski, Slawomir, and Clare Bates Congdon. "Realm: A rule-based evolutionary computation agent that learns to play mario." Computational Intelligence and Games (CIG), 2010 IEEE Symposium on. IEEE, pp. 83-90, 2010.
- [5] Ambinder, Mike. "Biofeedback in gameplay: How valve measures physiology to enhance gaming experience." game developers conference. Vol. 2011. 2011.
- [6] 棟方渚. "バイオフィードバックゲームを体験するユーザーの行動や印象." 研究報告エンタテインメントコンピューティング (EC) 2009.1, pp. 1-4, 2009.
- [7] Rani, Pramila, Nilanjan Sarkar, and Changchun Liu. "Maintaining optimal challenge in computer games through real-time physiological feedback." Proceedings of the 11th international conference on human computer interaction. Vol. 58. 2005.
- [8] 代蔵巧, 棟方渚, and 小野哲雄. "E3-Player: 鑑賞者の興奮を促進させる動画鑑賞システム." エンタテインメントコンピューティングシンポジウム 2013 論文集 2013, pp. 272-277, 2013.
- [9] J.Togelius,G.N.Yannakakis,S.Karakovskiy,andN.Shaker, "Believable Bots: Can Computers Play Like People?," Berlin, Heidelberg: Springer Berlin Heidelberg, ch. Assessing Believability, pp. 215-230, 2012.
- [10] Camilleri, Elizabeth, Georgios N. Yannakakis, and Alexiei Dingli. "Platformer level design for player believability.", Computational Intelligence and Games (CIG), 2016 IEEE Conference on. IEEE, pp. 1-8, 2016.
- [11] J. Togelius, S. Karakovskiy, and R. Baumgarten, "The 2009 Mario AI competition," in Evolutionary Computation (CEC), 2010 IEEE Congress on. IEEE, pp. 1-8, 2010.