

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 2 日現在

機関番号：34504

研究種目：挑戦的萌芽研究

研究期間：2013～2015

課題番号：25540172

研究課題名(和文)生物の基本原則の導入による「人間的で自然な強く無さ」を伴うゲームAIの自己組織化

研究課題名(英文)Autonomous Generation of Human-like Behaviors of Video-Game Agents using Biological Constraints

研究代表者

片寄 晴弘 (KATAYOSE, HARUHIRO)

関西学院大学・理工学部・教授

研究者番号：70294303

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：2000年後半以降、ビデオゲームのNPCの振る舞いの自動構成を目指した研究が行われるようになり、人間をはるかに凌駕する強いNPCが構成されるに至ったが、今度は、如何にして「人間的で自然な強くない」振る舞いを実装するかが課題として浮かび上がってきた。この課題に対し、ここでは、生物の「身体的な制約」と「生き延びるために必要な欲求」を強化学習、経路探索アルゴリズムに組み込み、アドホックにコーディングすることなく、NPCの「人間らしい」振る舞いを創出させる研究に取り組んだ。評価実験の結果、提案手法により、人間のプレイヤーより人間的と感じられるNPCの振る舞いが生成されること確認された。

研究成果の概要(英文)：Designing the behavioral patterns of video game agents (NPC) is a crucial aspect in developing video games. While various systems that have aimed at automatically acquiring behavioral patterns have been proposed and some have successfully obtained stronger patterns than human players, those patterns have looked mechanical. This study proposes the autonomous acquisition of NPCs' behaviors, which emulate the behaviors of human players. Instead of implementing straightforward heuristics, the behaviors are acquired using techniques of reinforcement learning with Q-Learning and pathfinding through an A* algorithm, where biological constraints; "sensory error", "perceptual and motion delay", "physical fatigue", and "balancing between repetition and novelty" are imposed. The subjective assessments using "Infinite Mario Bros" indicated that agents with biological constraints could acquire more human-like behavioral patterns than humans and agents without them.

研究分野：メディア情報処理

キーワード：ゲームプログラミング

1. 研究開始当初の背景

ゲームは、AI (人工知能) の絶好の研究対象として、長い歴史を持って取り組まれてきた。本研究を開始した当初、チェスについては、コンピュータが既に人間を凌駕しており、数年のうちに、人間のプロ棋士とほぼ互角のレベルとなることが見込まれていた。また、アクションゲームのゲーム AI のコンペティションにおいては、人間のプレイヤーを圧倒的に凌駕するパフォーマンスを発揮する NPC (Non Player Character) が自律的に構成できることが示された。このように、「強い」ゲーム AI の構成に関する研究は大きく進展した一方で、「人間らしい強くなさ」、「プレイヤレベルにあわせた自然な振る舞い」をどう構成していくのかについては十分に研究が進んでおらず、ゲーム AI の主要研究テーマの一つとして認識されるようになってきた。

2. 研究の目的

Computer が対戦相手、もしくは、パートナーを務めるビデオゲームのプレイ感の質を大きく左右するものとして、NPC の振る舞いのデザインは極めて重要である。これらのうち、「人間らしい強くなさ」、「プレイヤレベルにあわせた自然な振る舞い」については、これまで、プログラマが明にコーディングするという形で実装されていた。本研究では、これら人間的と判断される振る舞いを人間が直接コーディングするのではなく、自律的にゲーム AI に獲得させる手法についての検討を実施する。

3. 研究の方法

NPC の振る舞いに限らず、ヒトが何を見て「自然な」、あるいは、「人間的」とみなすかは、哲学的な視点から見ても極めて難しい課題である。その課題に対し、本研究では、物理的、あるいは、生物学的な制約下で、適応し、獲得した行動や思考パターンに対して、ヒトが「自然な」、あるいは、「人間的」などのラベル付けを行っているという仮説のもと、「生物学的な制約」を強化学習や探索アルゴリズム等に組み込んでいくことで問題の解決を図る。

3.1 生物学的制約

人間プレイヤーがゲームをするときには必ず、「ゆらぎ」「遅れ」「疲れ」といった生物学的制約が生じているはずであり、生物学的制約下で操作されたキャラクタの振る舞いは、人間にとって最も一般的で見慣れた振る舞いであると考えられる。本研究においては、具体的に、以下を「生物学的制約」として取り入れる。

1) センサ系、運動系における「ゆらぎ」

人間プレイヤーは、操作対象や敵オブジェクトなどの位置 (座標) を正確に観測し認識することは難しく、必ず誤差 (ゆらぎ) が生じる (見間違い、操作ミスなど)。そこで、ゲー

ム AI が観測する操作対象の現在位置やゲームの局面情報に対し、ガウスノイズを付与することで再現する。

2) 知覚から運動制御に至る「遅れ」

人間プレイヤーは、ゲームの局面を認識してから、実際に動作するまでに遅れが発生する (眼と手の協応動作における遅延など)。そこで、ゲーム AI が観測する操作対象の現在位置やゲームの局面情報を、数百ミリ秒過去の情報にすることで再現する。

3) キー操作の「疲れ」

人間プレイヤーは、ゲームのコントローラのキー操作を、きわめて短時間で何度も、または、長時間連続して実施すると疲れが生じる (ボタン連打、単調な操作の繰返しなど)。そこで、振る舞いを学習する際に、ゲーム AI にキー操作変更による負の報酬を与えることで再現する。

4) 訓練と挑戦のバランス

人間プレイヤーは、同じ行動を繰返すことで「訓練」する一方で、同じ行動の結果に飽きたり、その行動で失敗を繰返したりすると、飽きや失敗を解消するための新奇な行動に「挑戦」する。そこで、失敗を繰返しているゲーム局面では、新奇な行動に挑戦する傾向を高め、逆に、失敗をほとんどしないゲーム局面では、同じ行動を繰返して訓練する傾向を高めることで再現する。これは、強化学習における「探索と知識利用のジレンマ」と同等の考え方であるが、人間プレイヤーの試行錯誤学習のなかでも生じる現象であり、本研究では「訓練と挑戦のバランス」として定義する。

3.2 実験対象

本研究では、世界的に有名な 2D 横スクロール型アクションゲームである“スーパーマリオワールド”を模したゲーム“*Infinite Mario Bros.*”を対象として、人間らしい振る舞いの自動獲得に関する実験を実施する。ゲーム画面例を図 1 に示す。



図 1 *Infinite Mario Bros.* の画面例

“*Infinite Mario Bros.*”では、1) ゲームの仕様やゲーム環境パラメータが公開されている。かつ、2) 「敵や穴を避けてできる限り先に進む」という明確な目標が設定できるため、機械学習によるエージェントの振る舞い獲得に適している。また、当該ゲームは人間型キャラクタを操作可能であるため、評価者

は人間らしい振る舞いを想起しやすい対象といえる。

3.3 使用アルゴリズム

本研究においては、強化学習として、Q 学習、探索アルゴリズムとしては A*アルゴリズムを用いて、「生物学的制約」の導入の効果について検証する。

3.3.1 Q 学習

Q 学習は、強化学習手法のなかでも、ゲーム内での形勢を報酬という形で直感的に設定でき、また、最適なルールの獲得として学習が進む点で、ゲームプログラマが利用しやすいというメリットがある。

$$\operatorname{argmax}_{at} Q(st, at) \quad (3.1)$$

ここで、 t はゲーム開始からの時刻、 st は時刻 t におけるゲーム局面、 at は時刻 t においてエージェントが選択する行動、 $Q(st, at)$ は局面 st と行動 at の組に対する、 Q 値と呼ばれる評価値である。つまり、Q 学習では、局面 st において Q 値が最も高くなる行動が最適であると出力される。また、エージェントの行動に応じて、以下の式で Q 値を更新することにより学習が可能となる。

$$Q(st, at) = (1 - \alpha)Q(st, at) + \alpha(r + \gamma \max_p Q(st+1, p)) \quad (3.2)$$

式 (3.2) において、 α は学習率と呼ばれる、 Q 値の更新において新たな報酬 r をどれだけ重視するかを示す値、 γ は割引率と呼ばれる 0 以上 1 以下の定数である。 r は局面 st において行動 at を選択したことによって得られる報酬である。エージェントの行動選択手法としては ϵ -greedy 法を用いる。

リアルタイム性のあるビデオゲームにおける各項目の扱い方として、時刻 t はフレーム単位であり、局面 st や行動 at が無数に設定できる場合は、学習が実時間で収束するようゲーム特徴を考慮した状態圧縮が必要である。また、報酬 r として、操作対象の進んだ距離、経過時間、局面が遷移する際の評価値（形勢）の増減などを与えることで、エージェントの振る舞いの自動獲得が可能となる。Q 学習への人間の生物学的制約の導入に関しては、式 (3.1)、式 (3.2) のすべての st について、 n フレーム過去（遅れ）のエージェントの位置座標にガウスノイズ（ゆらぎ）を付与したゲーム局面 $st-n$ とすることで、「ゆらぎ」と「遅れ」を実現する。「疲れ」は、式 (3.2) の Q 値の更新の際に、報酬 r にキー操作変更による負の報酬を与えることで実現する。「訓練と挑戦のバランス」は、ランダム行動選択確率 ϵ の設定において、失敗を繰り返しているゲーム局面 st では大きな値を設定することで、新奇な行動に挑戦する傾向を高め、逆に、失敗をほとんどしないゲーム局面 st では小さな値を設定し、同じ行動を繰り返して訓練する傾向を高めることで実現する。

3.3.2 A*アルゴリズム

A*アルゴリズムは、ゲーム AI の行動選択においてはよく利用される経路探索アルゴ

リズムであり、以下の式によりゲーム木の経路のコストを算出する。

$$f^*(n) = g^*(n) + h^*(n) \quad (3.3)$$

ここで、 $g^*(n)$ は、スタートノードから、あるノード n を経由してゴールノードに辿り着くまでの経路の推定コスト、 $h^*(n)$ はヒューリスティック関数と呼ばれ、現在のノード n からゴールノードまでのコストの推定値である。生物学的制約の導入に関して、「ゆらぎ」と「遅れ」は、数百ミリ秒過去のキャラクターの位置情報に対してガウスノイズを付与し、その座標をスタートノードとすることで実現する。「疲れ」は、極めて短時間でのキー操作の変更を禁止することで再現する。「訓練と挑戦のバランス」は、学習フェーズを持たない A*アルゴリズムでは実現不可能であるため対象外とする。

4. 研究成果

4.1 生物学的制約の効果の検証

人間の生物学的制約を導入した Q 学習エージェント及び A*エージェントによって獲得された振る舞いが、本当に人間らしいかどうかを検証するため、20~24 歳の男女 20 名（男性 13 名、女性 7 名）を対象として評価実験を実施した。

実験にあたって、Q 学習においては、生物学的制約なし、すべての生物学的制約あり、「訓練と挑戦のバランス」のみ制約ありの三条件、A*エージェントについては、「ゆらぎ」「遅れ」「疲れ」にかかる制約がある場合とない場合の二条件、さらに、人間のプレイヤーとして、初級者、中級者、上級者のプレイヤーによる計 7 個の 10 秒程度のプレイ動画を用意し、これらに対して、シエッフエの一対比較法によって「人間らしさ」にかかる評価を行うという手続きを採用した。得られた結果を相対的嗜好度に基づいてプロットしたものを図 2 に示す。

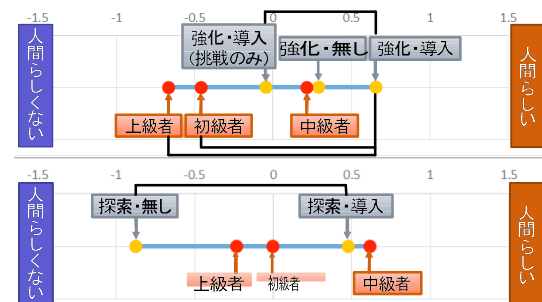


図 2 「人間らしさ」に関する相対的嗜好度 (上) Q 学習 (下) A*アルゴリズム

この図より、まず、Q 学習においてすべての生物学的制約を導入したプレイ動画が、人間のプレイヤーによる動画も含めた中で、最も「人間らしい」と評価されていることが読み取られる。

Q 学習エージェント内での比較結果については、人間の生物学的制約を導入した[強化,

導入] (相対的嗜好度: 0.66) は, 人間の生物学的制約を導入していない[強化, 無し] (相対的嗜好度: 0.29) と比較して, 人間らしいという結果が得られた. しかしながら, 相対的嗜好度の差については, 5%水準の有意差は認められなかった.

Q 学習エージェントと人間プレイヤーの比較結果については, 人間の生物学的制約を導入した[強化, 導入] は, 人間プレイヤーの[初級者][中級者][上級者] より人間らしいという結果が得られた. ただし, 有意差が認められたのは, [強化, 導入] と[初級者], [強化, 導入] と[上級者]であり, [強化, 導入] と[中級者]では有意差が認められなかった.

A*エージェント内での比較では, 生物学的制約を導入した[探索, 導入] は, 生物学的制約を導入していない[探索, 無し] と比較して, 1%水準で有意に人間らしいという結果が得られた. 最後に, エージェントと人間プレイヤーの比較結果を述べる. 生物学的制約を導入した Q 学習エージェント[強化, 導入]は, 人間プレイヤーの[初級者][中級者][上級者] より人間らしいという結果が得られた. また, 生物学的制約を導入した A*エージェント[探索, 導入] は, 人間プレイヤーの[初級者][上級者] より人間らしいという結果も得られた. ただし, 有意差が認められたのは, [強化, 導入] と[初級者], [強化, 導入] と[上級者], [探索, 導入] と[上級者] のみであった.

Q 学習において「生物学的制約」を導入したエージェントが生成した振る舞いに対しては, 「たじろいで見える」, 「ジャンプにおいて必要以上に高くみえる」等, 情緒に訴えるコメントが寄せられた. この例が示すように, 「人間らしい」と映る振る舞いを明示的に実装したわけではないが, エージェントの構成に「生物学的制約」を導入することにより, ヒトが人間らしいと判断する振る舞いが生成できることが示された.

本研究の過程や結果について国内外の主要会議で発表した所, 情報処理学会ゲームプログラミングワークショップ 2013, 国際会議 International Conference on Advances in Computer Entertainment Technology において, それぞれ, 最優秀論文賞, ACE Gold Paper Award を受賞し, 情報処理学会ゲーム情報学研究会から業績 1) の論文に関する招待を受けることになった.

4.2 被験者の評価基準に関する初期的検討

4.1 で述べたように, 研究は順調に進捗した. その一方で, 図 2 において, [強化, 導入] と[強化, 無し], [強化, 導入] と[中級者] の間に有意差が認められないなど, 実験参加者によって評価基準が大きく異なる可能性があることが浮き彫りとなってきた. この状況の把握を目的として, 任天堂のアクションゲーム「スーパーマリオオマージュアイランド」を対象とし, 操作に関する各スキルの難易度を定量化し, その状況に応じてスキルレベルを定量化した動画に対する評価実験に着手

した. 動画については, RTA (ステージを最速でクリアすることを目指したリアルタイムアタック) の日本記録であるプレイ動画を最高レベルとし, 「ジャンプができる」を最低レベルとした合計 9 つのものを用意した. 評定者には, 「評定の結果は実験の途中で何度変更しても良い」と教示し, 「上手さ」そのプレイが「人間のものであると考える確信度」を評定させた.

この結果, 「上手さ」については, 各スキルの積み上げによって評定されるレベルが, 主観的な視点でほぼ評定されること, 一方で, 中位のレベルの評定においては, 評定者間で若干のばらつきがあること, 「人間のものであると考える確信度」については, 上級レベルのものほど下がるということが確認された. また, 評定者の被験者レベルの分析からは, 中級者ほど「評定の結果を変える」回数が多いことが確認された. 4.1 で述べた評価実験は一対比較法を用いたためその懸念には当たらないが, 主観評価を必要とする研究においては評価者選定や実験デザインに細心の注意が必要であることが示唆された.

5. 主な発表論文等

[雑誌論文] (計 1 件)

1) 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」振る舞いの自動獲得, 情報処理学会論文誌, Vol. 55 No. 7, pp. 1655-1664, 2014. 7

[学会発表] (計 10 件)

2) Nobuto Fujii, Yuichi Sato, Hironori Wakama, Koji Kazai, Haruhiro Katayose: Evaluating Human-like Behaviors of Video-Game Agents Autonomously Acquired with Biological Constraints, Intl. Conf. on Advances in Computer Entertainment Technology 2013 (ACE2013), LNCS 8253, pp. 61-76, University of Twente Netherlands, 2013. 11

3) 藤井叙人, 福嶋良平, 片寄晴弘: エンタテインメント系システムの主観評価実験におけるユーザ統制及び実験手法の検討, 情報処理学会 エンタテインメントコンピューティング 2015, pp. 429-434, 札幌市教育文化会館, 2015. 9

[その他]

ホームページ

<http://ist.ksc.kwansei.ac.jp/~katayose/>

6. 研究組織

(1) 研究代表者

片寄 晴弘 (KATAYOSE HARUHIRO)

関西学院大学・理工学部・教授

研究者番号: 70294303