

## 2016年度 博士研究員研究成果報告書

氏名 (所属研究室) 藤井 叙人 (理工学研究科片寄研究室)

研究 課 題 エンタテインメントコンピューティングのための

ゲーム AI の構成に関する研究

研究 期 間 2016年4月1日～2017年3月31日

研究 成 果 概 要 (日本文 (全角) の場合は2,500字程度、英文 (半角) の場合は90字×65行程度)

本研究では、人間プレイヤーを『楽しませる』ゲーム AI を自動的に獲得するための方法論の構築、及び、その主観評価実験において信頼性を確保するための実験方針の策定を目的とした。

従来、ゲーム AI の自動獲得は、機械学習や経路探索の手法を用いて、ゲーム AI の戦略・戦術を獲得することで実現されてきた。ゲーム情報学領域ではコンピュータ将棋やコンピュータ囲碁において多大な成果があげられており、コンピュータがトッププロ棋士に完勝したというニュースは世界中の話題になったといえる。しかし、従来研究におけるゲーム AI の戦略・戦術は、過度に最適化されているため、不自然すぎる、あるいは、強すぎるものであり、人間プレイヤーにとっては機械的に映ってしまう。

ゲーム AI に対する要求事項を、エンタテインメントコンピューティング領域の観点から検討すると、ゲームをプレイした際のユーザエクスペリエンスをより良いものとし、人間プレイヤーのゲームへのモチベーションを維持することといえる。この事項を満たすゲーム AI の自動獲得は、従来研究のような、機械学習や経路探索を用いてゲーム AI の戦略・戦術を獲得する手法では実現が難しい。この課題を解決すべく、申請者はこれまで、十分に強くなったゲーム AI に「人間らしさ (自然さ)」を持たせる研究を実施してきた。人間プレイヤーがゲームをする際の生得的制約をモデル化し、モデルを導入した教師なし学習、及び、経路探索手法により人間らしいゲーム AI の自動獲得手法を提案した。また、獲得されたゲーム AI の振る舞いについて、人間的と感じさせる要因の主観評価/自由記述調査を実施し、「人間らしい」と感じさせるゲーム AI が生成できていることを確認した。

本研究期間では、その次のステップとして、ゲーム AI と一緒に遊ぶことによる『楽しさ』を計測するための指針策定を目標に、人間プレイヤーによる主観評価実験を実施してきた。主観評価実験における『人間らしさ』や『楽しさ』といった評価項目の定量化は、本研究のみならず、エンタテインメント系システム全般において、避けては通れない本質的な評価である。しかし、これらの項目は、かなり個人性が強い対象であるため実験参加者の統制が極めて難しい。そのような状況を踏まえ、EC 系システムの評価実験において考慮していくべき事項、特に、実験参加者の経験・知識の把握や統制に関して議論をする必要がある。そこで、本研究では、安定した評価実験を行なうにあたって考慮すべき統制視点の検討を実施した。実験参加者の経験や知識の量が実験結果に影響を及ぼしている可能性を指摘し、再現性のある主観評価実験を実施するにあたって考慮すべき観点を洗い出し、それに則して実験計画を立てることで、どの程度評価結果の信頼性が向上するのか検証した。

本研究の主観評価実験では、まず、「結果がぶれる原因となる、実験計画に帰する要因」として、実験参加者の熟練度と知識量、実験刺激の提示順序、実験参加者の評価基準と評価軸の3

つを挙げた。その上で、これらの要因が実験結果にどのような影響を与えるかを検証した。主観評価実験の対象として、1995年に発売されたアクションゲーム「スーパーマリオ ヨッシーアイランド」(国内売上本数は2013年時点で177万本の有名タイトル)を使用し、20~24歳の男女8名(男性7名、女性1名)を対象に実験を執り行った。

まずは、実験参加者の熟練度を知るために、家庭用ゲームの経験に関するアンケートを実施した。次に、実験参加者にヨッシーアイランドの操作方法を説明した後に、実際にゲームをプレイさせた。その際、実験参加者のヨッシーアイランドに対する知識や技術の量を知るために、録画録音したプレイ動画を、実験後に実験者がチェックした。チェックの際には、コントローラ操作の難しさにより4段階にレベル分けされた全45項目のチェックシートを用いている。最後に、複数のプレイ動画を実験参加者に見てもらい、うまいプレイと感じる順、及び、人間らしいプレイと感じる順に順位付けをさせた。実験参加者には「各動画は過去に遡って何回見ても良い」、「評定の結果は実験の途中で何度変更しても良い」と教示しており、さらに、評定中は、実験参加者には評定の理由について発話するように促し、その発話内容を録音した。

実験の結果、本ゲームの未経験者は、「うまさ」と「人間らしさ」について、順位付け(相対評価)はできても採点(絶対評価)するのは難しいことが示された。実験参加者の熟練度と知識量の違いによる「順位付けはできるが、点差を正確につけることができない」という状況が生じている可能性がある。また、本ゲームの未経験者、及び、初級者レベルの実験参加者は、実験の最中に評定結果を変更することが多々あり、これは、自分の想定外の上手さの動画であったために他を下げざるを得ない状況であることが分かった。実験刺激の提示順序の影響で、評定するための評価基準が大きく変更してしまった例と言える。評定中の実験参加者の発話を分析すると、実験中は一貫した評価軸により評定している実験参加者がいる一方で、本ゲームの上級者レベルの実験参加者の中には、実験の途中で明らかに別の評価軸を用いて評定している人がいることが分かった。これは、上級者になればなるほど、評価軸が複数ある可能性が高いと言える。

実験結果から、エンタテインメント系システムの主観評価の実験計画で注意すべき点をまとめる。まず、システムが対象とするエンタテインメントの未経験者を実験参加者とした場合、順位付け(相対評価)はできても採点(絶対評価)するのは難しいと思われる。絶対評価をさせたいのであれば、少なくとも何度かは経験したことのある参加者を用意するべきである。ただし、経験が浅い参加者は、実験途中で自身の想定以上(あるいは想定以下)の刺激に直面し、評価の基準(閾値)が実験過程で随時変更される危険性がある。上級者、熟達者ほどの経験がある参加者であれば、評定結果がそれほどぶれることなく、安心して採点等の絶対評価を任せられそうである。しかし、上級者になればなるほど、評価軸が複数ある可能性が高いという問題もある。あるコンテキストで考えると良い評価だが、そうでない場合は悪い評価であり、どちらの評価をすべきか悩むケースも起こりうる。

研究の成果は、情報処理学会エンタテインメントコンピューティング2016において、特別企画「EC研究を論文にするには」内にて口頭発表を実施した。また、情報処理学会誌EC特集号「豊かな体験をつくるエンタテインメントコンピューティング技術」において、解説論文「エンタテインメントの評価と脳科学」の一部に本研究の成果を記述している[情報処理学会誌, Vol.58, No.1, pp.16-19, 2017].