

Title	強化学習を用いた良好で多様なゲームコンテンツ生成
Author(s)	NAM, SANGGYU
Citation	
Issue Date	2022-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/18132
Rights	
Description	Supervisor:池田 心, 先端科学技術研究科, 博士

氏 名	NAM, Sanggyu		
学 位 の 種 類	博士（情報科学）		
学 位 記 番 号	博情第 476 号		
学 位 授 与 年 月 日	令和 4 年 9 月 22 日		
論 文 題 目	Using Reinforcement Learning to Generate Game Content with Quality and Diversity		
論 文 審 査 委 員	主査 池田 心	北陸先端科学技術大学院大学	教授
	飯田 弘之	同	教授
	長谷川 忍	同	教授
	白井 清昭	同	准教授
	Ruch Thawonmas	立命館大学	教授

論文の内容の要旨

Procedural content generation (PCG) has arisen as one major research field in games. PCG aims to provide content that meets specific criteria of quality, such as playability, difficulty, and entertainment. In addition, the generated game content should be different from those currently offered. Therefore, achieving quality and diversity becomes one of PCG's most important objectives.

Machine learning is a representative algorithm for PCG and is called procedural content generation via machine learning (PCGML). PCGML usually requires data for learning. However, collecting sufficient and good data can be costly. Another typical method is search-based PCG, where content is generated by evolutionary and other metaheuristic search algorithms. The search-based PCG does not require training data. Instead, it tries to optimize the given evaluation functions. The optimization process is time-consuming, making search-based PCG challenging to generate content immediately on demand.

To solve these problems, we apply reinforcement learning (RL), which does not require training data but learns from interactions with the environment. Another advantage of RL is that once the training is done, it can generate content quickly when required. In this dissertation, we target two kinds of games with distinct characteristics, turn-based role playing games (turn-based RPG) and Super Mario Bros. (Super Mario).

For turn-based RPG, we train RL agents to generate stages, where a stage is a series of events such as battles and recovery. For Super Mario, we train RL agents to generate levels, where a level consists of tiles such as walls and enemies. The generation is a challenging task since components in the stages and levels, such as the events and tiles, are highly correlated.

In order to address this challenge using RL, it is necessary to formulate the problem by a Markov decision process (MDP). Thus, we formulate the stage/level generation into MDPs, where the ideas can be generalized to other PCG problems. For the rewards in MDP, hand-crafted evaluation functions which reflect players' enjoyment are used to evaluate generated content.

For the stage generation problem, two classical RL algorithms are adopted. One is a deep Q-network (DQN) for discrete action spaces, and the other is a deep deterministic policy gradient (DDPG) for continuous action spaces. Experiments show that both DQN and DDPG can generate good stages, where those by DDPG generally receive higher scores from the evaluation function.

Next, we try to apply similar approaches to the level generation for Super Mario. However, Super Mario's level generation has several distinct challenges that do not exist in the stage generation. One is related to how to represent the levels. Unlike events in turn-based RPG that can be straightforwardly represented by a small number of parameters, tiles in Super Mario levels need to be represented by a large number of parameters. For this challenge, we employ conditional

generative adversarial networks (CGAN), which have succeeded in generating images. Another challenge is related to how to evaluate the levels' difficulty. For this challenge, we employ a human-like AI agent and have it play the generated levels. We use twin delayed DDPG (TD3) for the level generation problem. As a result, the generated levels receive high evaluation values indicating good quality.

Finally, virtual simulations that give rewards to intermediate actions are employed to obtain content with even higher quality. In addition, we introduce noise to avoid generating similar stages/levels while trying to keep the quality high. The experimental results highlight that the proposed methods successfully generate good and diverse stages/levels for turn-based RPG and Super Mario.

Keywords: Reinforcement learning, Procedural content generation, Turn-based RPG, Super Mario, Quality-Diversity

論文審査の結果の要旨

ゲームは人工知能技術の良い適用対象であり、機械学習、最適化、木探索など多くの手法がゲームを通じて提案されてきた。深層学習の登場によりコンピュータプレイヤーの強さを求める研究がおおいに進展し、最近では強さだけでなく人間プレイヤーを楽しませるための研究も注目されるようになっている。ゲームのマップなどを自動作成する PCG (Procedural Content Generation) はなかでも重要な領域であり、(1)既存のコンテンツを訓練データとして機械学習する PCGML や、(2)評価関数を定めてそれを最適化する PCGEA が主に用いられてきた。しかし、(1)には訓練データが大量に必要なこと、(2)では生成時のコストが高いことが問題とされてきた。

このような背景のもと Nam くんは、コンテンツ生成問題をマルコフ決定過程に定式化し、強化学習を用いて生成方策を学習するという極めて新規性の高いアプローチを提案した。具体的には、作成途中のコンテンツを状態、それを一步完成に近づける作業を行動、最終的なコンテンツの評価値を報酬として定式化し、深層強化学習を用いて学習することで、与えられた未完成コンテンツを評価値が高くなるように完成に近づける作業が高速に選べるようになる。この枠組み PCGRL は、事前の学習コストが高い一方で、訓練データを必要とせず、生成時のコストが低くなるため、PCGML および PCGEA と共存する PCG の新しい一つの柱となりうる。

さらに Nam くんは、この枠組みにおいて強化学習の Credit Assignment 問題を解決して学習の質を高めるための Virtual Simulation (VS)、質に配慮して得られるコンテンツの多様性を高めるための Diversity-Aware Greedy Policy (DAGP) という手法を提案し、実験によりその有効性を確かめている。具体的には、ターン制ロールプレイングゲームおよび、より複雑なゲームとしてスーパーマリオブラザーズのステージ生成にこの枠組みを適用し、いくつかの既存手法や、VS, DAGP を用いない場合との性能比較を行った。スーパーマリオブラザーズへの適用については、強化学習が行動として出力するステージ断片が高次元であること、ステージ評価には人間らしいコンピュータテストプレイヤーが必要なことなどのいくつかの課題があった。これに対して Nam くんは、CGAN を用いて小さい次元の出力から高次元のステージ断片に変換する手法や、人間の生物学的制約を考慮することで人間がミスしやすい箇所を推定する手法を用いて、高品質なステージ生成に成功した。

これらの結果はゲーム情報学のトップジャーナル IEEE ToG に採録済、および投稿予定である。

以上、本論文は自動コンテンツ生成の新しい有力な枠組みを打ち立てたものであり、学術的に貢献するところが大きい。よって博士(情報科学)の学位論文として十分価値あるものと認めた。