

Title	強化学習を用いたターン制RPGの多様なステージ自動生成
Author(s)	ナム, サンギユ
Citation	
Issue Date	2019-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/16720
Rights	
Description	Supervisor:池田 心, 先端科学技術研究科, 修士(情報科学)

Generation of Diverse Stages in Turn-Based Role-Playing Game using Reinforcement Learning

1610424 Nam Sanggyu

The artificial intelligence (AI) research mainly about the game has focused on developing strong AI player in diverse games. Although there are still many works to do, we can say that these are achieved by many cases such as Deep blue in chess, AlphaGo in Go, AlphaStar in Starcraft, OpenAI Five in dota2.

With their success and dramatic growth of AI, new challenges have arisen like developing teaching AI or entertaining AI. Currently, AI has been used effectively in various applications. One example involves the automatic generation of content using some algorithms in games, which is referred to as procedural content generation (PCG), and it is one of the major research fields in the game area. PCG can be applied to any game. PCG was mainly investigated by the generation of stages in popular games such as scroll action games (e.g., Super Mario brothers), racing games, real time strategy games (e.g. starcraft), and problems in puzzle games (e.g. puyo-puyo). On the other hand, relatively few studies reported turn-based RPG (RPG), which is a well-known classic genre.

In a majority of the RPGs, players play character roles to complete the story (e.g., defeat the boss), which often requires the growth of characters. Players can make their characters stronger by receiving rewards from successive battles. If players attempt to defeat all enemies using all their resources during successive battles, then it might be difficult to defeat the boss, or if players ignore all enemies, then characters may not have sufficient strength to defeat the boss. Hence, players must decide their own strategy, such as “Win this battle safely using some resources,” “Save mana or items to prepare for difficult battles,” or “Retreat against difficult enemies,” leading to the entertainment of the RPG.

To make available the various strategies, it is crucial to assign appropriate locations of enemies, frequency of recovery, and several parameters such as the strength of the enemies, the effectiveness of the items, and to the extent of degree to which the characters are cured, and others, which is the balance of the “Stage.” In addition, in terms of fun, it is crucial to provide diverse stages, which makes players decide different strategies. With these stages, RPG can avoid the monotony, and players can feel refreshed. Hence, the research goal involves the generation of these diverse stages, leading to entertainment.

PCG has several approaches. One example is procedural content generation via machine learning (PCGML). PCGML was actively investigated in

recent years. In several PCGML studies, network models use the existing game content data to learn generation. When considerable content created by human designers can be easily obtained, it might be possible to generate game content based on data distribution using generative models such as variational autoencoders (VAE), PixelCNN, or generative adversarial network (GAN). However, with respect to PCG, it is difficult to collect sufficient content used for training data. Furthermore, it may be not guaranteed that the generated content is desired by designers, and it may not be varied as it follows the distribution of training data. Hence, generate-and-test algorithms such as search-based PCG are typically used. The training data are not required in search-based PCG, but the human-defined evaluation function is needed instead, which exhibits an advantage in that each evaluation function can be tailored to the direction required by the designers or by specific players according to their skills or taste. Typically, generate-and-test algorithms generate content by the optimized evaluation or fitness value using a genetic algorithm (GA). However, GA is slightly slow in instantly providing content to various players; in addition, GA possibly generates “Similar Content Group” in principle.

Thus, reinforcement learning (RL) is considered to solve this disadvantage. Assume that the complete stage comprises n discrete sections, “State” is any stage with n or less sections, and “Action” is the generation of the next section, which is not generated thus far. When the stage is completed, then its good or bad is evaluated by the designed evaluation function, and its value is given as a “Reward.” With the progress of learning, a desired stage defined by the evaluation function is generated.

Two RL models, Deep Q-Network(DQN) and Deep Deterministic Policy Gradient(DDPG), respectively, are selected, and the generated stages are evaluated as 0.78 and 0.85 by our designed function, respectively. We knew that it is possible to generate “good stages” by using reinforcement learning, but just good stages are not enough, “qualified and diverse stages” are also important. So, we proposed the method for selecting the “noised second-best generation which is not lower quality and somewhat differs from the best generation” rather than the “highly evaluated best generation”

We call this policy as the Stochastic noise policy (SNP), and by the application of the SNP, diverse stages are successfully obtained, and those diversities are evaluated by the parameter mse and the different number of valid strategies. As a result, generated stages scored the average evaluation value as 0.82 which means good enough and parameter mse was 0.52 and valid strategies was also diverse. So we verified the various and good stages can be obtained