アカデミックライティング調査レポート「草稿」

**ゲームAIと機械学習・強化学習**

2EP1-59　宮澤　瑛至

平成31年5月14日提出

**あらまし：**本レポートでは，ゲームAIと人工知能の違いや，これらに導入されつつある機械学習の1つの強化学習について文献調査を行い，それに基づいた計算リソースの課題解決のためのシステムについての考察を述べている．

1. **まえがき**

　今日のビデオゲームの飛躍的な発展において，ゲーム内のキャラクター制御，いわゆる「ゲームAI」に，より人間らしさを与え，よりゲームへの没入感を与えるため，機械学習の1つである強化学習が導入され始めている．本レポートでは，これらゲームAIと人工知能の違いや発展，機械学習の導入について文献調査を行い，その結果に基づいて，強化学習導入の課題とそれを解決するシステムから，今後の学習計算リソースについて考察する．

以下，2章からゲームAIにおいての人工知能について（文献[1]に基づく），3章でゲーム内への機械学習の応用（文献[2]に基づく），4章で学習計算リソースに関する展開，5章でまとめと感想を述べる．

1. **ゲームAIにおいての人工知能**
   1. **人工知能とゲームAIの違い**

　ゲームにおける人工知能というのは，コンピュータの動かすキャラクターであるノンプレイヤーキャラクター（以下，NPCと呼ぶ）が，知能を持って動いているように錯覚させるための動きのプログラムのことを指す．これを「ゲームAI」と呼ぶ．プレイヤーの心地よさや，人間らしい動きを目指したものであるため， 実際の人工知能とは異なる動きをする．

　従来のゲームAIで制御されるNPCの動きとしては，キャラクターの監視を常に行い，それに応じてNPCに指示をするような，俯瞰的なプログラムを中心とした考え方がほとんどであった．しかし，ゲームのスケールや複雑さが従来よりも増し，1体のNPCが処理する情報量や計算量が上昇ため，キャラクター自身が周囲の状況を認識し，自ら意思決定を行って動くように，NPCを中心とする自律的なプログラムの考え方に転換されていった．こうした考え方を，「自立型AI」または「自立型エージェント」と呼ばれる[1]（図1を参照）．

※青円内はゲーム空間

監視・指令

自立型AI

ゲーム空間を認識し,

意思決定

図1 従来と自立型の比較 (文献[1]を基に筆者作成)

従来のゲームAI

* 1. **ゲーム内NPCの意思決定**

ゲーム内NPCの意思決定におけるアルゴリズムとして，代表的なものが7つ存在する．

**表1 意思決定アルゴリズムの種類 （文献[1]をもとに著者作成）**

|  |  |  |
| --- | --- | --- |
| ルールベース | ルールを列挙し，選択して実行し制御する | |
| ステートベース | 動作の状態とその推移から制御する | |
| ビヘイビアベース | 状態をツリー方式にし，振る舞いや行動により制御を変える | |
| タスクベース | 問題解決のためのタスクを順序に従って行う | |
| ゴールベース | | 常に目標を設定し，行動に指向を持たせて制御する |
| ユーティリティベース | | 行動に対する効用を上がるように制御する |
| シミュレーションベース | | 自身の行動を推測させて制御する |

この7つのアルゴリズムには，長所と短所があり，得意とするゲームシステムが存在するため，開発されるゲームシステムによって，こうしたアルゴリズムを選択，あるいは組み合わせてAIを制作することがほとんどである．

1. **ゲーム内への機械学習の応用**

近年の人工知能ブームにより，本来人工知能と働きが異なるゲームAIでも，人工知能の技術が取り入れられている．その多くが，人工知能の意思決定に使われる機械学習を応用している．機械学習により，プレイヤーや自身の行動を学習して，より人間らしい動きをさせるのが目的である．

* 1. **機械学習の原理**

機械学習とは，大量のデータから規則性を見つけ出し，推論のためのルールを機械に生成させる手法である．大量のデータの特徴を抽出し，推論のためのルール（モデル）を生成した後，改めて与えられたデータと生成した推論モデルを照合して，データがどのような種類であるかを判断する．データから特徴を集めた推論モデルを生成することを「学習」といい，与えられたデータをモデルによって判断することを「推論」という．機械学習には，「教師あり」または「教師なし」の学習と強化学習に分けられる。

　強化学習については，ある環境下においたうえで行動し，それに応じた報酬が最大値になるように方策を求める手法である．本レポートで取り上げているゲームAIや自動運転，ロボット制御のシミュレーションなどに利用されている．

　こうした機械学習の手法のほかにも，模範学習やカリキュラム学習などがある[2]．

* 1. **強化学習の原理**

　ゲームAIにおいて強化学習は，NPCの行動を最適化させることが目的で行われる．強化学習の特徴は環境に学習し行動するキャラクター（以下，エージェント）が，行動に対して得られる報酬が最大化されるような推論モデルを生成することである．[2]

　エージェントは，環境下で行動を起こし，もしそれが環境に対して望ましい結果であれば正の報酬，そうでなければ負の報酬を受け取り，次により多くの報酬を得られる行動を選択する．こうした，行動，評価，学習のサイクルを繰り返すことで，エージェントの行動を最適化していく（図2参照）．

環境

エージェント→

右

状態・報酬

行動

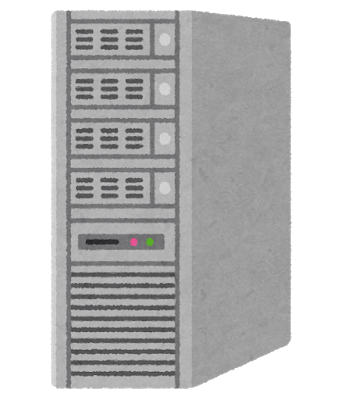
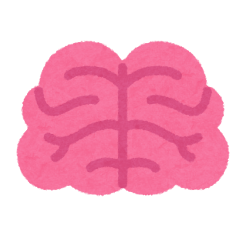
図2 強化学習のサイクル （筆者作成）

* 1. **ゲームAIにおける機械学習での課題**

　ゲームAIの開発において，ゲームが複雑化し，より多くの環境情報量が多くなると，AIの学習する時間が長くなり，ゲームの進行に影響する[3]．また，計算リソースであるメモリーも，現状，グラフィックや物理演算，音声データなどで多くが占有されていることがほとんどであり，リソースの限られるスマートフォンなどではこうしたAIの実装が難しい場合がある．

1. **学習計算リソースに関する展開**

学習するための計算値ソース確保が課題となっているが，私はそれらをクラウド技術によって解決できると考えられる．実際に提案され，研究が進んでいるが，実用した例は未だ無い．そこで自分なりに問題を解決できるシステムを考えた．

サーバーと実機をネットワークにより接続し，サーバー上では学習させるAIの脳に当たる推論モデルデータなどを位置させ，実機上ではプレイヤーと学習によって実際に動くエージェントキャラクターを置く．ゲームが進行している間，エージェントが得た環境下での状態をサーバーに送信する．サーバーでは受け取った状態を元に，評価，学習をした後に次に行う行動を決定して実機に送る．

ゲーム内

エージェント

学習

次の行動

状態

図3 システムのイメージ図 （筆者作成）

サーバーで学習処理を肩代わりすることにより，実機のリソースを圧迫することなくAIを学習させることが可能である．ゲームはリアルタイムな対応が求められるため，通信速度によって遅延が生じる可能性があるが，第五世代通信技術（5G）が実用化に近づいているため，今後リアルタイムにゲームAIを学習でき，近年実用化が進むと考えられる．

1. **むすび**

　本レポートでは，第2章に，ゲームAIと人工知能の原理的な違いについてと意思決定の手法について述べ，第3章では，人工知能技術の応用である機械学習および強化学習についての原理と課題，そして第4章に，その課題とする学習計算リソースの問題をクラウドによって解決する方法を考察した．今後，人工知能の発達により，より人間らしく，面白さのあるゲームAIが誕生することを期待したい．

**文献**

[1] 三宅 陽一郎 著，“知性を表現する手法”，“人工知能の作り方～「おもしろい」ゲームAIは以下にして動くのか～”，秋山 絵美（技術評論社）（編），2017年1月5日 初版，pp.59

[2] 布留川 英一 著，“Unity ML-Agentsの全体像”，“Unity ML-Agents 実践ゲームプログラミング”，佐藤 英一（編），2018年7月25日 初版，pp.15~pp.19

[3] 甲野 佑 著，“強化学習技術とゲーム AI 〜 今できる事と今後できて欲しい事 〜”， [https://www．slideshare．net/yukono1/ai-82646808](https://www.slideshare.net/yukono1/ai-82646808) ，Published on Nov 24，2017，pp.40