



成長するノンプレイヤーキャラクターの知能とその転移

東京工芸大学 工学部

助教 河野 仁

1. はじめに

近年、TV ゲームやコンピュータゲームなどのデジタルゲームにおいて、AI 活用が活発である。ゲームキャラクター（エージェント）単体の AI として実装されるのではなく、ゲーム全体に制御可能なメタ AI やナビゲーション AI など提案され、ユーザによるデジタルゲームの楽しみ方が多様化している。そのため計算機の発展や AI 技術の発達により、ユーザをより楽しませるだけでなく、開発者側の作業負担にも貢献できると考えられる。今後は機械学習などの技術をゲームのエージェントに応用することが重要であると考えられるが、教師データを用いせずに自分で解を探索可能な強化学習はエージェントの知能化に有用なアルゴリズムである。しかし、ニューラルネットワークや強化学習など代表的な学習アルゴリズムにおいても実装上の課題がある。一般的に、学習アルゴリズムは長時間の学習が必要であり、学習フェーズが時間的ボトルネックとなる。そのため、知能ロボットや知能化システムの研究領域では、過去に獲得した知識を再利用する転移学習という手法が提案されている。これらの基礎技術をエージェントに応用すれば、持続的なゲーム AI の成長や開発者作業負担軽減も可能であると考えられる。そこで、本研究で強化学習と転移学習技術を活用し、ゲームのエージェントがゲームタイトルを超えて学習結果を継承、さらにさまざまなタスク（ゲーム内での仕事）を経験することで、エージェント（NPC）の振る舞いを多様化できるようにするシステムの開発を目的とする。これにより、将来的にはエージェントのバックグラウンドで動作する AI が、演劇や TV 番組の役者に位置づけられるような役割を実現しプレイヤーは、ゲーム性やキャラクターの外見だけでなく、エージェントの「中の人」すなわちゲーム AI をも楽しめるシステムの実現を目指している。

2. 活性化拡散モデルを用いた方策選択手法

強化学習は、環境から与えられる報酬を基に、多くの報酬を獲得できる方策を学習するアルゴリズムであり、転移学習は獲得した知識を再利用する手法である。しかし、これらの技術はある 1 つの仕事しか学習が出来ず、多様な振る舞いを実現することはで

きない。そこで、本研究では認知心理学で提案されている活性化拡散モデルを参考に、獲得済みの複数方策を適応的に選択し、転移学習を行う手法を開発した。

提案手法は、エージェントが観測する環境情報を基に、ネットワーク構造上に保存された方策を活性化し、候補となる方策を選択する手法である。特徴としては、活性化した方策の隣接する方策も同時に活性化し、選択候補を複数用意する点である。活性化を行動毎に繰り返すことで、使用方策を先読みする効果が期待できる。

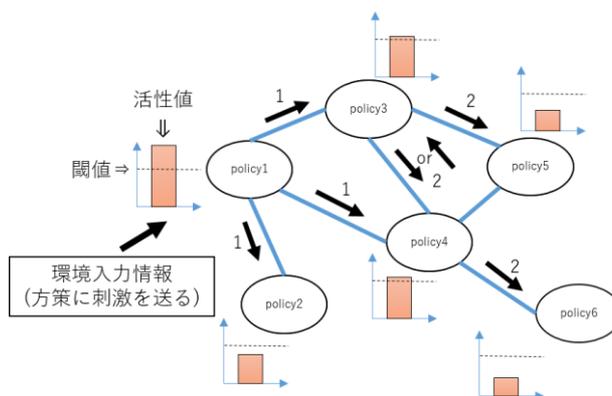


図1 活性化拡散モデルに基づく方策選択手法

3. 提案手法の効果検証

評価実験では、エージェントが過去に獲得した方策群から環境に効果的な方策を選択できるかを確認する。学習するタスクは、最短経路問題という迷路探索のシミュレーションを採用し、環境は図2に示すような環境を用いる。これによりエージェントが行動可能な範囲、行動が不可能な範囲である障害物、エージェントの目的地であるゴールの3種類のグリッドで構成する。エージェントは前後左右に1マスずつ1回の行動で移動することが出来る。学習器にはQ学習を用いる。

知識選択を行うための方策は、100種類用意しその中には再利用に効果的な方策と、そうでない方策が混在している。本実験では、それらの方策の中から図2の最短経路を解くのに有益な方策が選択できるか10回実験を行った。

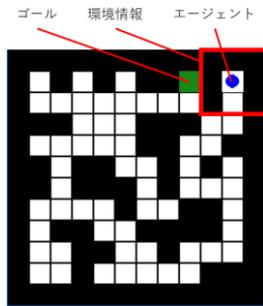


図2 実験環境

実験によって得られた結果を図6に示す。図3は学習曲線であり、横軸は学習の繰り返し回数、縦軸は目的を達成するまでに必要であった行動回数を表す。赤線は、方策の再利用を行わない通常の強化学習(Q学習)であり、この学習曲線が基準となる。黒線はあえて、負の転移が発現するように効果的でない方策を転移した場合の学習曲線であり、緑線は逆に効果的であることが判明している方策を転移した場合の学習曲線である。その他の色の学習曲線は、提案手法の学習曲線であるが、他の曲線と比べても値が低く曲線の収束スピードが早い。方策の再利用に効果的な知識を多く選択していることも実験結果から確認している。

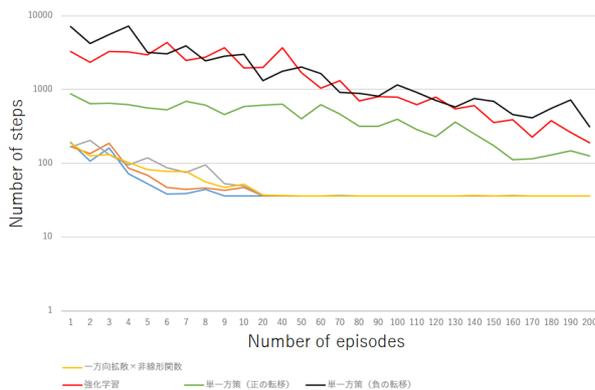


図3 提案手法を用いた実験結果

4. ゲームAIへの提案手法の実装

提案手法を実際のデジタルゲームシチュエーションで評価するために、図4に示す環境構築を行った。ゲーム要素を司るソフトウェアはUnityで構築し、ゲーム内エージェントの制御はC#にて開発した。また、学習機能はC#のエージェント制御プログラムとTCP/IP通信にて連携するPythonのプログラムを開発した。提案手法をNPCに実装するためには、あらかじめNPCに強化学習で獲得した複数の方策を持たせる必要がある。そこで、図5に示すような、ランダムに出現するプレイヤーエージェントをゾンビ(NPC)が捕獲する訓練シミュレータを構築した。また、提案手法をNPCに実装し、実際のプレイヤーとのインタラクションを評価するために、ゾンビから逃避するゲーム(図6)や、武

装したキャラクタと闘いながら逃走するゲームを開発した。しかし、TCP/IP通信が学習速度のボトルネックとなり、学習に長い時間が必要となることが明らかとなった。そのため、現在はNPCの訓練シミュレータによる予備的な学習が終了しておらず、実際のプレイヤーを交えた提案手法の評価には至っていない。継続して、システム開発と学習、評価を行っていく。

5. まとめ

本研究では、NPCが様々なゲームにおいて多様な振る舞いを実現し、さらにはゲームタイトル間を超えたゲームAIの開発を目的として、認知心理学の知見を活用した強化学習における転移学習の方策選択手法を提案した。実験結果から、提案手法を実装したエージェントが効果的な方策を選択しながら迷路の最短経路を探索し、転移学習の効果も発現した。また、提案手法を評価するためにUnityを用いたゲーム開発を行った。学習プログラムとキャラクタ制御プログラム間のTCP/IP通信がボトルネックとなり、NPCに使用させる方策の獲得に時間を要した。そのため、提案手法のユーザも交えたゲーム評価は今後の課題であり、継続して研究開発を行っていく。

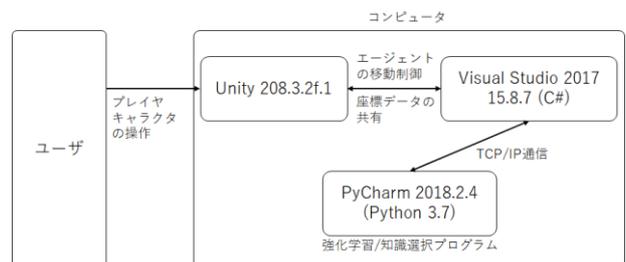


図4 ゲームAIの開発環境



図5 方策取得のための予備的強化学習



図6 開発したゲーム例(ゾンビからの逃避ゲーム)