

2011 年度 修士論文要旨

ビデオゲームエージェントの自律的行動獲得と観測情報の信頼性に着目した獲得行動パターンの分析

関西学院大学大学院 理工学研究科
情報科学専攻 片寄研究室 佐藤 祐一

近年、ハードウェアの発展によるハードウェアやグラフィックなどのリアリティの向上とともに、COM プレイヤ（ビデオゲームエージェント）の振る舞いや戦略が重要視されており、エージェントの行動パターン実装の自動化が課題となっている。その課題に際し、強化学習によってエージェントの行動パターンを自律的に獲得するアプローチの提案がなされているが、過度の最適化が進められる結果、強さが状況に応じて著しく揺らいだり、また、獲得された行動パターンが「機械的」と感じられるという問題があった。本研究ではその課題に際し、人間のもつ制約であるセンサ系や制御系の「ゆらぎ」「遅れ」といった要素を強化学習の学習系に組み込むことによって、人間がプレイしているような感覚の提供ができているかどうかの検証を行う。

本研究では研究達成のため、次のような課題がある。1 つ目は「ゆらぎ」や「遅れ」といった制約のビデオゲームエージェントへの組み込みである。これを「観測情報の信頼性」としてパラメータ化し、観測情報を変化させることによって制約をもたせた。2 つ目は観測情報が変化することによる強化学習の性能のゆらぎである。強化学習は状況の変化に弱いため、観測情報が変化することにより性能に大きなゆらぎが発生してしまう。3 つ目は強化学習アルゴリズムの決定である。ビデオゲームにおいては、環境が完全な既知ではなく、また報酬分布のモデルが与えられていない。ビデオゲームではある状態においてとった行動に対して直接的な評価をしなければならない。

本研究では Infinite Mario Bros を学習対象とし、3 つを課題を解決する。観測情報の信頼性パラメータによる観測情報の変化に対応するため、真の観測情報を予測することで解決した。また、学習方法としては環境や報酬分布の知識が必要なく、行動を直接評価できる Q 学習を用いて学習を行った。検証実験の結果、エージェントに対して「安全性」を持たせることができ、状況の変化に対応できることが明らかになった。また、被験者を対象に視聴実験による初期的検討を行った結果、「ゆらぎ」「遅れ」といった要素を強化学習に組み込むことにより「人間らしい」と感じる有意傾向がでていることが明らかになった。