

## システム情報工学研究科修士論文概要

年 度	平成 27 年度	学位名		修士( 工学 )
専 攻	知能機能システム	専攻	著者氏名	作田 宏行
指導教員氏名	安信 誠二			
論文題目	位相変化量の動的な更新による複素強化学習の学習性能向上に関する研究			
論文概要	<p>強化学習とは、エージェントと環境が相互作用することにより、試行錯誤を繰り返して状態に応じた望ましい行動を選択するエージェントを構築する学習制御の枠組みである。強化学習では、不完全知覚問題が課題となる。不完全知覚問題とは、学習器の知覚が不十分であるため、異なった状態が同じ観測として知覚され、正しい行動決定ができない問題である。不完全知覚問題を解決する手法として、行動や状態の観測情報の履歴を用いて学習を行う複素強化学習が提案されている。複素強化学習は、複素数平面上で定義された価値をステップ毎に位相回転することによって更新するため、不完全知覚の生じる状態でも行動選択確率が変更され、それぞれ異なる行動を選択可能なエージェントを構築できる。複素強化学習では、学習が進むにつれて、良い行動と悪い行動における複素行動価値が分離されていき、複素行動価値を内部参照値と比較することによって良い行動が選ばれやすくなる。しかし、一度分離された複素行動価値について、分離が阻害される現象が観察される。</p> <p>そこで、学習を決定する要素である位相変化量を動的に更新することで、従来では学習に時間を要していた不完全知覚を含む環境下で学習時間を短縮することのできる複素強化学習手法を提案する。提案手法では、従来の位相変化量に内部参照値と直前の複素行動価値との位相差を一定分乗じた値を加えて動的位相変化量を求める。動的位相変化量の導入によって、内部参照値が直前の複素行動価値に近づくため、複素行動価値の分離が阻害されにくくなる。</p> <p>まず、複素行動価値を固定して実験を行い、位相変化量が動的に更新されていることを確認した。次に、1種類の不完全知覚を含む環境下で実験を行った結果、提案手法によって学習が進んでいることを確認した。最後に、従来では学習に時間を要していた複数の不完全知覚が繰り返し現れる環境下で実験を行った。その結果、学習の終盤において、従来手法では学習途中で一旦収束したステップ数が再び増加してしまっていたが、提案手法では収束したままになった。これによって、従来では学習に時間を要していた不完全知覚を含む環境下で学習時間を短縮でき、性能が向上したことを確認した。</p>			
審査日	平成 28 年 1 月 28 日			
審査員	(大学名 職名)	(学位)	(氏名)	
主査	筑波大学 教授	工学博士	安信 誠二	
副査	筑波大学 助教	博士(工学)	澁谷 長史	
副査	筑波大学 教授	Ph.D.(工学)	堀 憲之	