

## 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	鎌田 晋太郎
指導教員氏名            安信 誠二				
論文題目  段階的知識拡充による多次元空間における行動知識の強化学習				
論文概要 <p>強化学習は機械学習の分野において相互作用に基づく目標指向型の学習(goal-directed learning)である。主な特徴として3点がある。(1)学習に際して教師信号を用いない。すなわち、環境のモデルを直接的に必要としない。(2)行動と評価のサイクルを繰り返しながら学習が進む。すなわち、動的な環境にもある程度は対応できる。(3)確率的な行動規則の獲得が可能のため、ノイズの多い実環境にもある程度対応できる。</p> <p>実環境では、状態が多次元なものが存在する。状態が多次元になると状態空間が大きくなり、強化学習では行動知識を獲得することが困難となる。その理由として、多次元空間ではエージェントが行動知識を獲得するためには膨大な回数の試行錯誤が必要となるからである。</p> <p>人は行動知識を獲得する場合、段階的に獲得している。自動車の運転技術習得を例に挙げる。教習生は自動車教習所にてアクセル・ブレーキの操作の習得から実際の道路での走行技術の習得まで段階的に習得する。本研究では、人が行っている段階的な行動知識獲得を学習に組み込むことで多次元状態空間において少ない試行回数で行動知識を獲得する手法を提案することを目的とする。</p> <p>本研究で提案した手法の有用性を7×7×7グリットワールド(3次元状態空間)で検証する。シミュレーション結果から従来の強化学習に比べて、提案する手法のほうが少ない試行回数で行動知識を獲得できた。さらに、最短経路の行動知識を学習していることも確認できた。本研究では、前のレベルの行動知識を知識拡充し、試行を繰り返すことによって現在の行動知識を獲得することを繰り返す手法を提案した。これにより、多次元状態空間において従来の強化学習に比べて少ない試行回数で知識を獲得した。</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 教授	Ph.D.(工学)	堀 憲之	
副査	筑波大学 助教	博士(工学)	澁谷 長史	