

## 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	大西 杜諒
指導教員氏名 安信 誠二			
論文題目			
支援対象者と障害物の相対角を用いた効率的な強化学習に関する研究			
論文概要			
<p>人は知覚することにより危険を察知し、障害物を回避するなどの行動を決定する。また人によって知覚できる方向範囲が異なる。特に高齢者などは目や耳などの機能が衰えたことにより危険の察知できない方向範囲が増加する。このような人の支援を行うためには個性にあわせた支援システムが必要である。強化学習はこのような支援システムを開発するための有力な手法の一つである。しかし、従来の強化学習で支援対象者を中心にした支援システムを構築した場合、支援する方向ごとに独立して支援行動を学習する必要がある。そのため、行動を学習するまでに多くの時間が掛かってしまうという問題が生じる。本論文では効率的な学習を行うため、支援が必要な方向範囲と必要でない方向範囲それぞれにおいて、支援対象者と障害物の相対角を用いた行動価値関数により高速に学習し、それらを合成して行動を決定する強化学習法を提案する。</p> <p>提案手法は効率的な学習を行うために「支援が必要な方向範囲での行動価値関数と必要でない方向範囲での行動価値関数を合成する手法」と、「相対角を用いた行動価値関数で高速で学習する手法」を組みわせ、それぞれの方向範囲で高速な学習を行うものである。高速化で学習する手法には原点をエージェントとした極座標システム内の支援対象者と障害物の相対角を用いる。そのため、方向による区別はできないが、学習が高速化される。相対角を用いた行動価値関数は二つの成分によって構成される。一つは距離と行動に依存する成分、もう一つは相対角に依存した成分である。価値の更新は距離と行動の成分のみ行えばよいため、学習の次元が低減される。二つの行動価値関数を合成する手法では二つの行動価値と方向のみを学習する選択関数を組み合わせ合成する。これらにより、それぞれの方向範囲において高速に学習することで、全体の効率化が図られる。</p> <p>提案手法の有効性を評価するために実験を行った。有効性の評価には従来手法として Sarsa と比較を行った。実験において提案手法は Sarsa と比べて短い試行回数で高い成功率を出すことができる結果を得た。これにより、それぞれの方向範囲において支援対象者と障害物の相対角を用いて学習した行動価値関を合成して行動を決定する強化学習法で効率的な学習を行えることを確認した。</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 准教授	博士(工学)	延原 肇