



強化学習における状態フィルタの適応的獲得手法に関する研究

永吉, 雅人

(Degree)

博士 (工学)

(Date of Degree)

2007-03-25

(Date of Publication)

2009-07-15

(Resource Type)

doctoral thesis

(Report Number)

甲4020

(URL)

<https://hdl.handle.net/20.500.14094/D1004020>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



【 3 6 7 】

氏 名・(本 籍)	永吉 雅人	(大阪府)
博士の専攻分野の名称	博士 (工学)	
学 位 記 番 号	博い第458号	
学位授与の 要 件	学位規則第5条第1項該当	
学位授与の 日 付	平成19年3月25日	

【 学位論文題目 】

強化学習における状態フィルタの適応的獲得手法に関する研究

審 査 委 員

主 査	教 授	上原 邦昭
	教 授	玉置 久
	教 授	小島 史男

本論文では、強化学習の実用化に向け、強化学習における状態空間の設計問題および不完全知覚問題の解決を目的として、状態フィルタの枠組みおよび実現手法を提案した。さらに、実際的な問題への適用例として、電動車いすの直進走行システムを取り上げ、電動車いす使用者の操作負担軽減に対する有効性について検討した。

第 2 章では、強化学習方式の概要の説明を行ない、現在強化学習方式が抱える多くの問題の中から盛んに研究がなされているものを取り上げ、それぞれの問題に対して代表的アプローチを示した。そして、本研究において状態空間の設計問題および不完全知覚問題に取り組むことを述べた。

第 3 章では、エージェントの観測空間と行動学習器の間に状態フィルタを導入した計算モデルを提案し、理想的状態フィルタについても記述した。つぎに、マルコフ決定過程 (Markov Decision Processes : MDPs) の場合と部分観測マルコフ決定過程 (Partially Observation MDPs : POMDPs) の場合について、状態フィルタを獲得する方法の違いという観点から従来手法の分類を行い、従来手法の構成面での特徴を明確にした。そして、内部状態を分割・統合する手法の状態フィルタの基本的機能として、内部状態の分割・統合方法について通常分割・統合・履歴参照分割・削除の 4 つの調整法を提案した。

第 4 章では、状態空間の設計問題に焦点をあて、タスクを遂行するために十分な情報が与えられている MDPs の問題を対象として、エントロピーを用いた状態フィルタの一実現法を提案した。この際、ある内部状態における行動選択確率のエントロピーを、その内部状態についての状態集約の正しさを評価する指標として用いた。この手法は、状態フィルタの学習と行動学習を同時並行して行なわせることができ、学習器において適用できる強化学習手法が限定されないという特徴をもっている。さらに、強化学習問題の例題として離散状態空間を有する迷路問題および連続状態空間を有するロボットナビゲーション問題を取り上げ、計算機実験を通して提案手法を従来手法・グリッド分割法と比べることで、提案手法の有効性・可能性を確認した。

第 5 章では、POMDPs への対応・状態空間のコンパクト化に焦点を当て、タスクを遂行するためには十分でない部分的な情報が与えられている POMDPs の問題を対象として、第 4 章で提案した手法を POMDPs に拡張する形で、適応的に履歴情報の記録・参照を行い、繰り返し内部状態を分割・統合することで POMDPs へ接近する状態フィルタの一実現法を提案した。この際、ある内部状態における行動選択確率のエントロピーを、その内部状態の細かさの十分性を評価する指標として用いた。この手法は、第 4 章において提案した

手法と同様、状態フィルタの学習と行動学習を同時並行して行なわせることができ、学習器において適用できる強化学習手法が特定されないという特徴をもっている。さらに、強化学習問題の例題として離散状態空間での迷路問題および連続状態空間でのロボットナビゲーション問題を取り上げ、計算機実験を通して、提案手法を従来手法と比べることで、提案手法の有効性・実問題への適用可能性を確認した。

第 6 章では、第 5 章で提案した手法のより実際的な問題への適用例として、電動車いすの直進走行システムを取り上げる。ここではまず、まちにあるバリアとして傾斜路面に注目し、三次元上における電動車いすの基本ダイナミクスのモデル化を行った。つぎに、基本的なダイナミクスを考慮した上、電動車いすの直進走行システムとして、第 5 章で提案した手法を再構成し、適用することによって、DC モータへの入力電圧を調整するシステムを提案した。この提案システムは、次のことが期待できる。1) 使用者一人一人に合わせた適切な制御ルールが実現されること、2) 未経験の路面を走行する場合や使用者の重心位置、車輪の空気圧が変化した場合など、使用者を含む電動車いすの特性が変化した場合においても、適切な制御ルールが維持されること、3) 使用者を含む電動車いすに関して、センシングする情報が部分的であっても適切な制御ルールが構築されること。そして、電動車いすの基本ダイナミクスに基づいたシミュレーションによって、(1) 傾斜角が一定な路面を走行する場合、(2) 傾斜角が一定でない路面を走行する場合、さらに補足として、(3) 重心位置が変化する場合、(4) 2 つの路面を走行する場合、について、提案システムの有効性ならびに実際的な適用可能性を確認した。また特に、提案システムを用いることによって使用者の入力負担を軽減できることも合わせて確認した。

以上より、本論文では、状態空間の設計問題と不完全知覚問題に焦点をあて、これらの問題が解消される有効な状態フィルタの実現手法を提案した。さらに、より実際的な問題への適用例として、電動車いすの直進走行システムを取り上げ、提案システムの有効性ならびに実際的な適用可能性を確認した。

氏名	永吉 雅人		
論文題目	強化学習における状態フィルタの適応的獲得手法に関する研究		
審査委員	区分	職名	氏名
	主査	教授	上原 邦昭
	副査	教授	小島 史男
	副査	教授	玉置 久
	副査		
			印
要 旨			
<p>本論文は、強化学習の実用化に向けて、強化学習における状態空間の設計問題と不完全知覚問題に焦点をあて、強化学習エージェントの状態空間を構成する状態フィルタの自動獲得・調整法の構築を目的とするものである。論文では、状態フィルタを導入した計算モデルを示すとともに、2つのクラスの問題に対する状態フィルタの適応的獲得手法を提案し、計算機実験を通してその有効性を確認している。さらに実際的な問題への適用例として、電動車いすの直進走行システムを取り上げ、電動車いす使用者の操作負担軽減に対する有効性を確認している。</p> <p>まず、第2章では、強化学習方式の概要の説明がなされており、現在強化学習方式が抱える多くの問題の中から盛んに研究がなされているものを取り上げ、その問題に対してそれぞれの代表的アプローチを整理している。そして、本研究において状態空間の設計問題および不完全知覚問題に焦点をあてることを示している。</p> <p>第3章では、エージェントの観測空間と行動学習器の間に状態フィルタを導入した計算モデルを提案し、理想的状態フィルタを定義している。つぎに、2つのクラスとして、マルコフ決定過程(Markov Decision Processes: MDPs)の場合と部分観測マルコフ決定過程(Partially Observation MDPs: POMDPs)の場合について、状態フィルタを獲得する方法の違いという観点から従来手法の分類がなされており、従来手法の構成面での特徴を明確にしている。さらに、内部状態を分割・統合する手法の状態フィルタの基本的機能として、内部状態の分割・統合方法について通常分割・統合・履歴参照分割・削除の4つの調整法を示している。</p> <p>第4章では、状態空間の設計問題に焦点をあて、タスクを遂行するために十分な情報が与えられているMDPsの問題を対象として、エントロピーを用いた状態フィルタの一実現法を提案している。そして、強化学習問題の例題として離散状態空間を有する迷路問題および連続状態空間を有するロボットナビゲーション問題を取り上げ、計算機実験を通して提案手法を従来手法・グリッド分割法と比べることで、提案手法の有効性・可能性を確認している。</p> <p>第5章では、POMDPsへの対応・状態空間のコンパクト化に焦点を当て、タスクを遂行するためには十分でない部分的な情報が与えられているPOMDPsの問題を対象として、第4章で提案した手法をPOMDPsに拡張する形で、適応的に履歴情報の記録・参照を行い、繰り返し内部状態を分割・統合することでPOMDPsへ接近する</p>			

氏名	永吉 雅人
<p>状態フィルタの一実現手法を提案している。そして、強化学習問題の例題として離散状態空間での迷路問題および連続状態空間でのロボットナビゲーション問題を取り上げ、計算機実験を通して、提案手法を従来手法と比べることで、提案手法の有効性・実問題への適用可能性を、それぞれ確認している。</p> <p>第6章では、まちにあるバリアとして傾斜路面に注目し、左右DCモータから駆動力を得る後輪駆動型電動車いすを対象として、まず三次元上における電動車いすの基本ダイナミクスのモデル化を行っている。つぎに、基本的なダイナミクスを考慮した上、電動車いすの直進走行システムとして、第5章で提案したPOMDPsにおける状態フィルタを用いた強化学習手法によって、DCモータへの入力電圧を調整するシステムを提案している。そして、電動車いすの基本ダイナミクスに基づいたシミュレーションによって、提案システムの有効性、実際的な適用可能性を、特に、提案システムを用いることによって使用者の入力負担を軽減できることを、それぞれ確認している。</p> <p>以上のように、本研究は、強化学習の実用化という観点から状態空間の設計問題と不完全知覚問題に焦点をあて、2つのクラスに対する状態フィルタの実現手法を提案するとともに実際的な問題への適用する場合のガイドラインを示すものであり、計算機知能の分野のみならず広くシステム工学・システム情報学の分野について重要な知見を得たものとして価値のある集積であると認める。よって、学位申請者の永吉雅人は、博士(工学)の学位を得る資格があると認める。</p>	