

Ansprechpartner:

Beginn: ab sofort

Dauer: 6 Monate



Philipp Karg, M. Sc.

IRS, Raum 206

Tel.: 0721/608-42708

philipp.karg@kit.edu

experimentell anwendungsorientiert theorieorientiert

Ihre Interessen:

Modellbildung Dynamische Optimierung
 Identifikation Statische Optimierung
 Reglerentwurf



Masterarbeit

Effizientes Sampling-basiertes Inverse Reinforcement Learning

Motivation:

Verfahren des Reinforcement Learning (RL) zielen darauf ab, die optimalen Reaktionen eines Agenten auf Zustandsänderungen in der Umwelt zu erlernen. Sie gehen dabei von einer bekannten Belohnungsstruktur aus und nutzen diese zur optimalen Adaption des Verhaltens des Agenten an seine Umgebung. In verschiedensten Anwendungen ist diese Belohnungsstruktur jedoch vorher unbekannt, d.h. es wird ein optimal agierender Agent und seine Aktionen beobachtet und Ziel ist es, die unbekannte Kostenfunktion dieses Agenten zu bestimmen.

Verfahren, die dieses inverse Problem lösen, werden als Inverse Reinforcement Learning (IRL) bezeichnet. Sie eignen sich z.B. zur Ermittlung der Belohnungsstruktur eines menschlichen Experten. Diese kann anschließend auf einen humanoiden Roboter übertragen werden (Learning by Demonstration) oder zum Design einer unterstützenden Automation in einer Mensch-Maschine-Kooperation genutzt werden. Aktuelle Verfahren benötigen hierfür jedoch meist eine große, nicht praxistaugliche, Anzahl an Beobachtungen (Samples) des Experten.



Aufgabenstellung:

Ziel der Arbeit ist die Entwicklung eines effizienten Sampling-basierten IRL-Verfahrens. Dabei soll ein im Rahmen einer Vorgängerarbeit entwickelter IRL-Algorithmus als Ausgangspunkt dienen. Die für das Verfahren benötigte Größe der sogenannten Zustandssumme, welche normalerweise aus den Expertensamples bestimmt wird, soll dabei über einen neuartigen Ansatz berechnet bzw. approximiert werden. Ansatzpunkte

können effiziente Sampling-Strategien basierend auf einem Modell der Umwelt sein oder ein Erlernen der Zustandssumme mittels Funktionsapproximatoren, wie Neuronale Netze. Der Nachweis der Anwendbarkeit der Methoden durch Simulationsbeispiele rundet das Aufgabenpaket schließlich ab.

