Andreas Dörr arbeitet im Bereich Künstliche Intelligenz an Maschinen, die autonom lernen neue Probleme zu lösen. In seiner Arbeit verbindet Dörr Erfahrungen aus der Luft- und Raumfahrttechnik im Bereich Systementwicklung und Regelungstechnik mit Methoden der künstlichen Intelligenz, des Machine Learnings, und der Robotik. Nach seinen Luft- und Raumfahrt- und Informatik Studien an der Universität Stuttgart promovierte er zu diesen Themen am Max-Planck-Institut für Intelligente Systeme und dem Bosch Center for Artificial Intelligence. Mit Forschungsaufenthalten am NASA Ames Research Center und dem Massachusetts Institute of Technology, legt seine Arbeit Grundsteine um die Methoden des maschinellen Lernens und speziell des Reinforcement Learnings aus simulierten, akademischen Umgebungen in reale Produkte und Prozesse zu übertragen. Als Forschungsingenieur am Bosch Center for Artificial Intelligence beschäftigt sich Andreas Dörr mit der automatisierten Regelung von komplexen Systemen, wie z. B. der Regelung von Fahrzeugverhalten in Extremsituationen.

Forschungsbeitrag zu Künstlicher Intelligenz in realen Systemen

Andeas Dörr

Prädiktive Modelle für daten-effizientes Reinforcement Learning

In aktuellen industriellen Anwendungen und Produkten begegnen uns eine immer größere Systemkomplexität und ein Zusammenwirken von vielen verschiedenen Teilsystemen. Gleichzeitig erwarten wir eine möglichst effiziente Nutzung aller Ressourcen bei optimaler Systemperformance. Bei der Regelung und Ansteuerung vieler dieser Systeme ist allerdings unklar, wie eine optimale Lösung aussehen könnte und gegebenenfalls steht der Aufwand um eine solche Lösung zu finden in keinem Verhältnis zum erwarteten Gewinn. Klassischerweise werden Regelungskonzepte für solche Systeme von Experten anhand von (physikalischem) Modellverständnis oder anwendungsspezifischen Heuristiken abgeleitet und analysiert. Die Modellierung dieser Systeme stößt jedoch oft an Grenzen, sei es durch die Komplexität der physikalischen Wirkmechanismen, die Größe und Komplexität des Problems oder Begrenzungen in Budget oder Zeit.

Aktuelle Methoden des *Maschine Learnings (ML)* und der *Künstlichen Intelligenz (KI)* eröffnen einen neuen und abweichenden Lösungsansatz. Dabei wird nicht mehr explizit Wissen über das System und den Weg zur möglichen Lösung vorausgesetzt, sondern die Maschine selbst muss Charakteristika des Problems und ihrer Umgebung erfassen und die optimale Lösung zur gestellten Aufgabe selbst finden. Vorwissen über das System kann zwar eingebracht werden, die wesentliche Charakterisierung erfolgt jedoch über eine abstrakte Beschreibung des gewünschten Verhaltens. Die Maschine selbst muss nun identifizieren, welche konkreten Abläufe erforderlich sind um die abstrakten Ziele zu erreichen.

Autonomes Lernen von neuem Verhalten

Beim *Reinforcement Learning (RL)*, oder auch bestärkendem Lernen, geht man idealisiert von einer Interaktion eines Agenten mit einer zunächst unbekannten Umgebung aus. Der Agent kann durch Aktionen Einfluss auf seine Umgebung ausüben und mittels seiner Sinne gewisse Auswirkungen seiner

Kurzzusammenfassung:

Eines der großen Versprechen von Machine Learning (ML) und Künstlicher Intelligenz (KI) ist es, einer Maschine die Fähigkeit zu verleihen, in komplexen und unbekannten Umgebungen autonom neues Verhalten zu erlernen. Ergebnisse dieser Forschungsrichtung, im speziellen der Disziplin des Reinforcement Learnings (RL), sind allerdings nach wie vor überwiegend beschränkt auf sehr spezialisierte, simulierbare und gut verstandene Problemfelder. Die Arbeit von Andreas Dörr adressiert dagegen Probleme wie: Daten-Effizienz, Robustheit und Lernen in Gegenwart von fehlerhafter und unvollständiger Wahrnehmung der Umgebung. Diese Probleme treten typischerweise auf, sobald maschinelle Systeme in unseren unstrukturierten, alltäglichen Umgebungen operieren sollen. Kernthemen seiner Arbeiten sind das Erlernen probabilistischer, prädiktiver Modelle, die es der Maschine erlauben aus vergangenen Erfahrungen optimale Entscheidungen für zukünftige Handlungen abzuleiten.

2019-WvSR-AusgezForschungsbeitraege_01-132_druck.indd 25 18.11.19 20:08

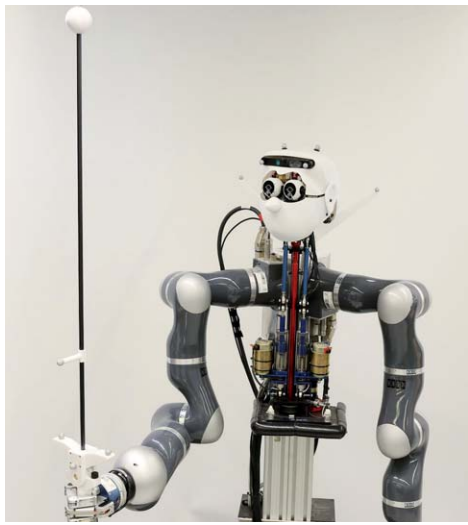

Abbildung 1:
Methoden des Reinforcement Learnings ermöglichen es dem Roboter Apollo das invertierte Pendel in seiner Hand zu balancieren. Im Gegensatz zu Fertigungsrobotern in Industrieanlagen hat er zunächst kein Wissen über sich und seine Umgebung. Er ist nicht vorprogrammiert und muss selbst herausfinden, wie er z. B. seinen Arm optimal ansteuern kann.

Aktionen auf die Umgebung beobachten. Gleichzeitig erhält der Agent von Zeit zu Zeit Rückmeldung, ob er sein Ziel erreicht oder erfüllt hat oder diesem zumindest näher gekommen ist. Diese Rückmeldung, auch Reward genannt, versucht der Agent langfristig zu maximieren. Diese Methode des bestärkenden Lernens, mit positiver Belohnung oder negativer Bestrafung bei erwünschtem oder unerwünschtem Verhalten ist aus dem Tierreich bestens bekannt und untersucht.

Mit Methoden des Maschinellen Lernens, können die allgemeinen Prinzipien des Reinforcement Learnings auch auf die eingangs erwähnten, automatisierte und maschinelle Systeme übertragen werden. Ein Beispiel ist der Roboter Apollo in *Abbildung 1*. Dieser Roboter bildet einen humanoiden Oberkörper mit den Freiheitsgraden zweier menschlicher Arme und Hände nach. In diesem Fall, ist der Agent der Computer bzw. das Computerprogramm, das die Motoren in den Gelenken des Roboters als „Aktionen“ ansteuern kann um mit der Umgebung zu interagieren. Über Sensoren in den Gelenken können gleichzeitig die aktuelle Position der Arme und über Kameras die Umgebung erfasst werden. Eine Aufgabe könnte es nun sein, den im *Abbildung 1* gezeigten Stab auf der Handfläche zu balancieren. Diese Aufgabe des invertierten Pendels ist ein klassisches Problem der Regelungstechnik. Die Physik kann einfach modelliert werden und eine Vielzahl an Lösungsansätzen existiert. In diesem Fall, hat der Agent bzw. der Computer jedoch kein Vorwissen über die Physik des Problems bzw. seine Umgebung und das invertierte Pendel. Lediglich die Beobachtungen aus der Interaktion zwischen Roboterarm und