

Einsatz von Reinforcement Learning in der Modellfahrzeugnavigation

von

Manuel Trittel

Informatik

HAW Hamburg

Vortrag im Rahmen der Veranstaltung AW1 im Masterstudiengang, 02.12.2008

Einführung

Einführung

- Thema
- Gesamtsystem der Anwendung
- Themeneinordnung
- Rahmenbedingungen

Reinforcement Learning

- Definition
- Begriffe
- Algorithmen

Methodisches Vorgehen

Risiken

Ausblick

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Einführung

Thema

Einsatz von Reinforcement Learning in der Modellfahrzeugnavigation

Konkreter Anwendungsfall:

Geschwindigkeitsregelung

**Geschwindigkeitsmaximierung
= Zeitminimierung**

Einhaltung einer maximalen Zentripetalkraft

Gliederung

Einführung

- **Thema**
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Einführung

Gesamtsystem

Videosequenz

Gliederung

Einführung

- Thema
- **Gesamtsystem**
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Einführung

Themeneinordnung

Aktuelle FAUST-Themen in der Modellfahrzeugnavigation:

Fahrspurerkennung (TFALDA)

Steuerung

Kartografie

Ausweichen

Einparken

Überholen



Gliederung

Einführung

- Thema
- Gesamtsystem
- **Einordnung**
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Einführung

Rahmenbedingungen

Vorhandene Regelungsalgorithmen

Messung der Zentripetalkraft
Beschleunigungssensoren
Gyroskop

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- **Rahmenbed.**

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

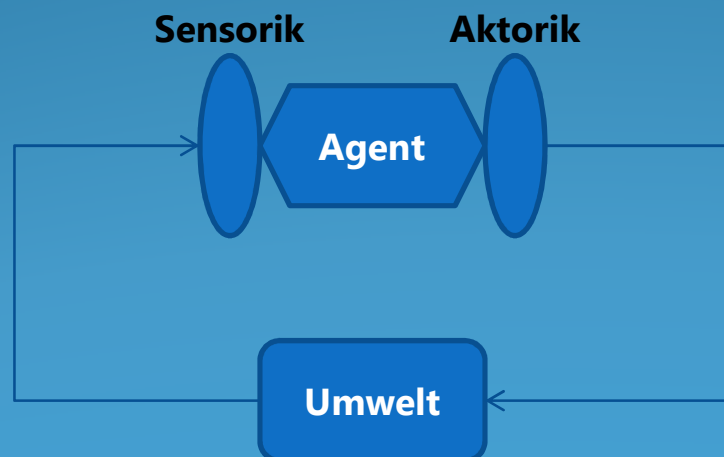
Reinforcement Learning

Definition

Lernsituationen beim Maschinellen Lernen

Überwachtes Lernen	(supervised learning)
Unüberwachtes Lernen	(unsupervised learning)
Bestärkendes Lernen	(reinforcement learning)

Trial-and-Error Prinzip



Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- **Definition**
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken


Ausblick

Reinforcement Learning

Begriffe

- **s State** **Zustand**
- **a Action** **Aktion**
- **r Reward** **Belohnung**

- **π Policy** **Strategie**

- **Optimale Strategie**  **Maximale Gesamt-Belohnung**

- **Markov Decision Process (MDP)**
 Grundlegende Theorie.
 Diskrete Zustandsraster mit Wahrscheinlichkeiten für
 Zustandsübergänge und Aktionen

- **Markov-Eigenschaft**
 Wahrscheinlichkeitsverteilung eines Zustands
 unabhängig von bisheriger Historie des Systems

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- **Begriffe**
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Reinforcement Learning

Algorithmen

Temporal Difference Learning (TD-Learning)

- Lernen ausschließlich durch Erfahrung (*Monte Carlo Methode*)
- Anpassung der Bewertungen auf Basis vorheriger Werte (*Dyn. Progr.*)

			25	50	25
			25	50	100
S		25	50	100	Z

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- **Algorithmen**

Method. Vorgehen

Risiken

Ausblick

Reinforcement Learning

Exploration vs. Exploitation

Was tun bei unbekannter Umwelt

Explore (dt. „erkunden“)

- Neue Wege ausprobieren

Exploit (dt. „ausbeuten“)

- Bisher Gelerntes optimal ausnutzen

Ansatz: ϵ – Greedy Suche

- Mit Wahrscheinlichkeit ϵ zufällige Aktion wählen
- Mit Wahrscheinlichkeit $1-\epsilon$ beste gelernte Aktion
- ϵ kontinuierlich verringern

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- **Algorithmen**

Method. Vorgehen

Risiken

Ausblick

Reinforcement Learning

TD-Learning Verfahren

Q-Learning

- Off-Policy-Verfahren
- Aktionsauswahl unabh. von zu optimierender Strategie

SARSA

- On-Policy-Verfahren
- Zu optimierende Strategie auch zur Aktionsauswahl
- Namensbildung aus $s_t, a_t, r_{t+1}, s_{t+1}, a_{t+1}$

• Lernraten

- Einfluss neuer Erfahrungen beim Update der Policy

• Diskontierung

- Einbeziehung möglicher zukünftiger Rewards
- „Weitsichtigkeit“ des Agenten

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe

• Algorithmen

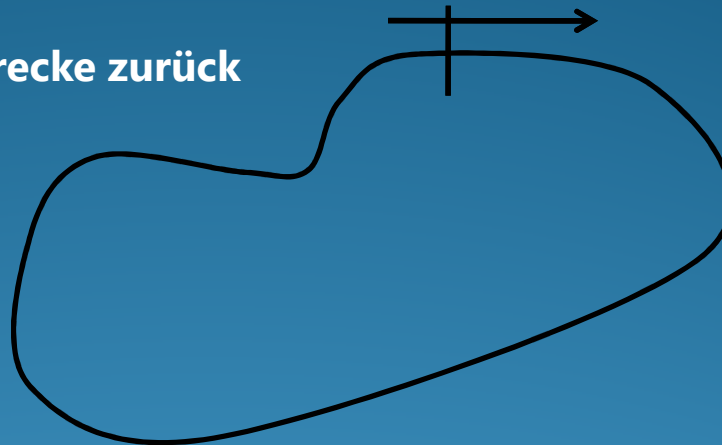
Method. Vorgehen

Risiken

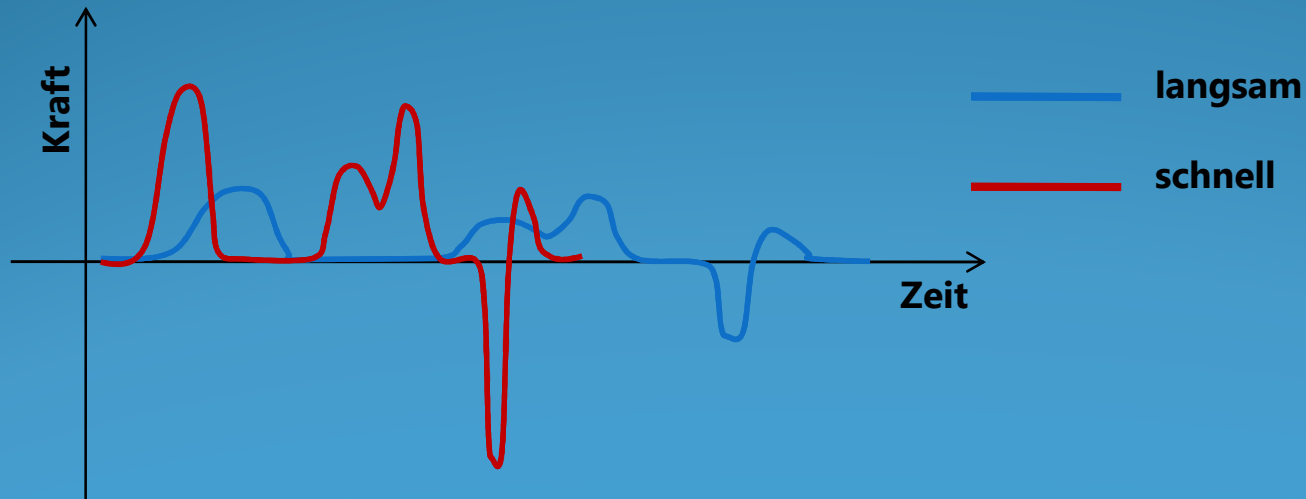
Ausblick

Methodisches Vorgehen

Fahrzeug legt eine beliebige Strecke zurück



Sensordaten bei konstanter Geschwindigkeit aufnehmen



Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

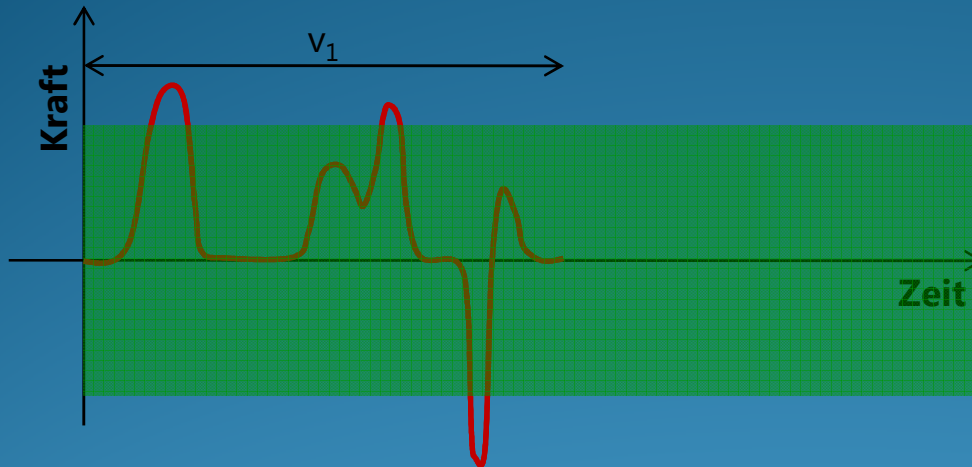
Method. Vorgehen

Risiken

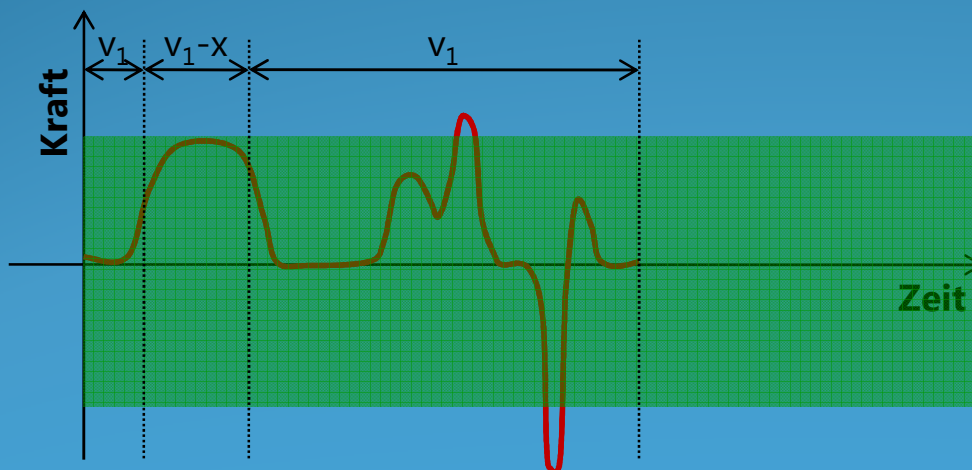
Ausblick

Methodisches Vorgehen

Festlegung einer maximalen Zentripetalkraft



Ausgleich kritischer Peaks



Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Methodisches Vorgehen



Funktionsapproximation mit Hilfe von

- Radialen Basisfunktionen (RBF)
- Kubischen Splines (KS)

Fahrzeug bremst ab

- Zentripetalkraft sinkt
- Benötigte Zeit steigt

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Methodisches Vorgehen

Umsetzung in der Praxis

Analyse der Eignung von RBF und KS

Erstellen einer Simulation

Portierung auf das Modellfahrzeug

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Risiken

Mögliche Problempunkte

Abweichungen Raum-Zeit Korrelation

- **Wo befinde ich mich auf dem Kurs?**
- **Korrekte Anpassung im Zeitbereich bei Ausgleich der Peaks**

Qualität der Sensordaten

- **Noch nie getestet (Einbau ins Fahrzeug folgt)**

Bestimmung der maximalen Zentripetalkräfte abhängig von

- **Bodenbeschaffenheit**
- **Bodenhaftung (Reifenqualität)**

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Ausblick

Weitere Anwendungsfälle

Vom Kleinen zum Großen

Selbst Erlernen der maximalen Zentripetalkraft

Lenkwinkelregelung für optimale Kurvendurchfahrt

Gliederung

Einführung

- Thema
- Gesamtsystem
- Einordnung
- Rahmenbed.

Reinf. Learning

- Definition
- Begriffe
- Algorithmen

Method. Vorgehen

Risiken

Ausblick

Vielen Dank für die Aufmerksamkeit!

Fragen?