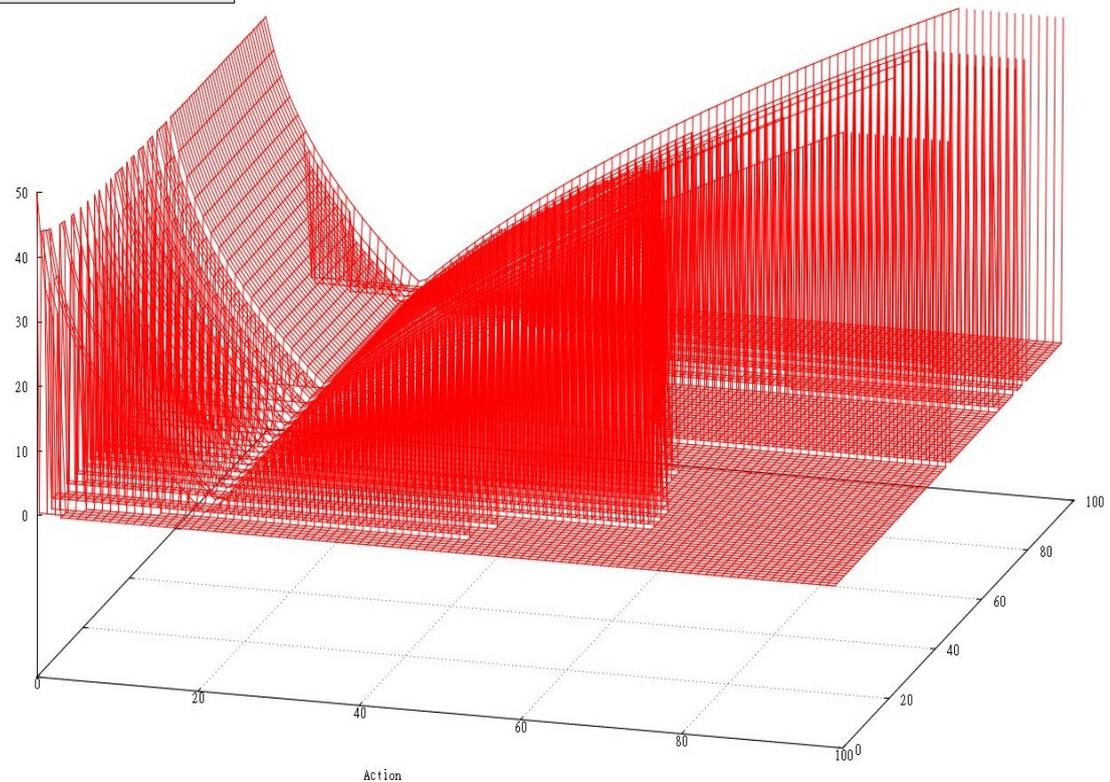


Lernen von Modellen im Reinforcement Learning

Jan Busch

Betreuer: Prof. Dr. Pareigis



Agenda

- Motivation
- Reinforcement Learning
- Hintergrund / FAUST
- Problemanalyse
- Idee / Risiko
- Ausblick

Seite 2

Motivation

- Wie lernt ein Mensch?

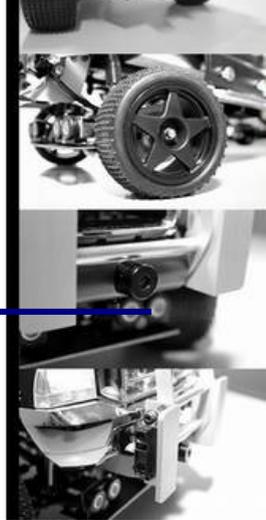


- Wie eine Maschine?



Seite 3

Motivation



- Erlernen einer optimalen Steuerung (z.B. FAUST)



- Beschleunigen und Verbessern von Lernvorgängen
- Lernen in „sicherer“ Umwelt → Simulation

Seite 4

FAUST



HAW HAMBURG

Fahrerassistenz und
autonome Systeme

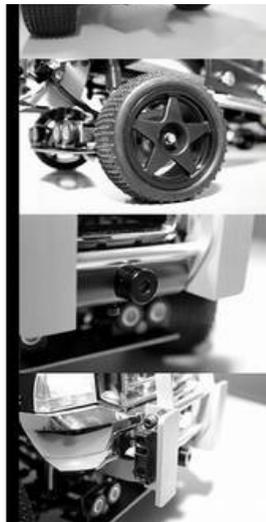
Sensorik

Steuerung

Bildverarbeitung

Algorithmen

Software

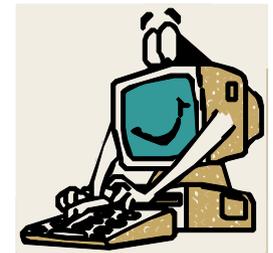


Seite 5



Reinforcement Learning

- Maschinelles Lernen → Bestärkendes Lernen
- Agent interagiert mit seiner Umwelt
- Positive bzw. negative Belohnung
- Modellfrei
- Trial & Error



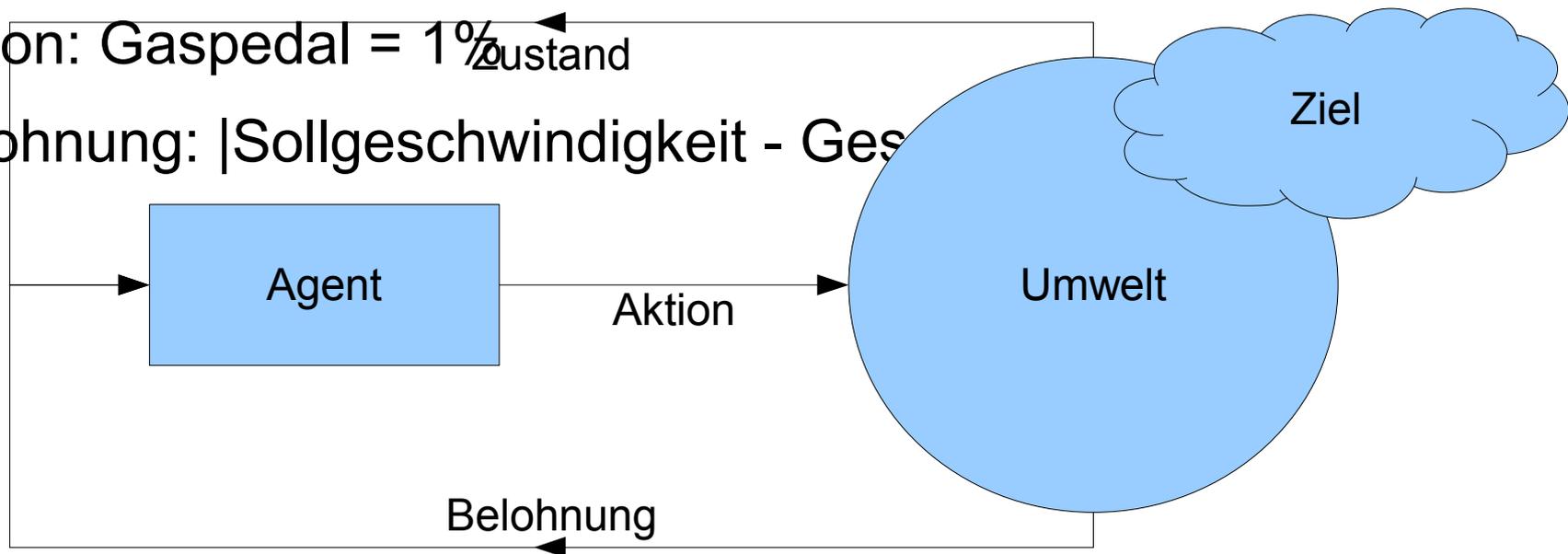
Reinforcement Learning

Ziel: Geschwindigkeit = 50 km/h

Zustand: Geschwindigkeit = 0 km/h

Aktion: Gaspedal = 1%

Belohnung: $|\text{Sollgeschwindigkeit} - \text{Ges}$



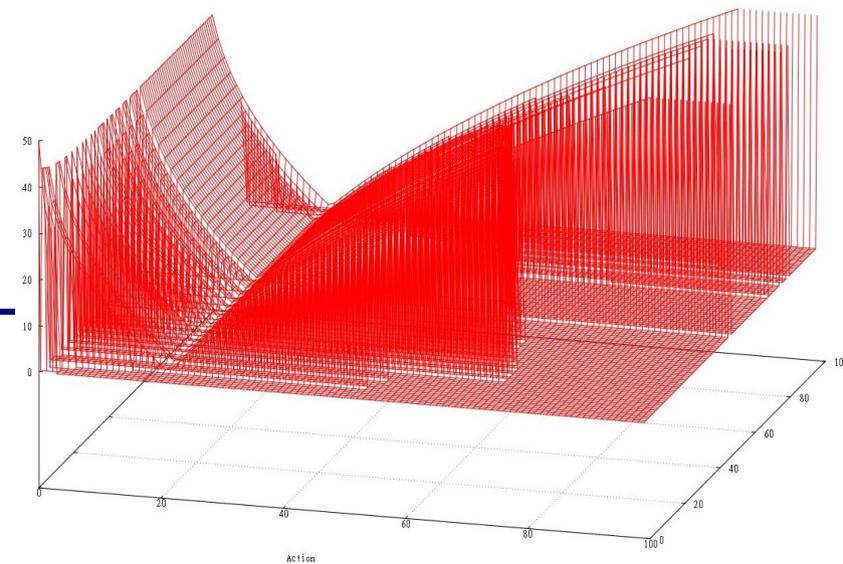
Das Problem

- Gelerntes nicht übertragbar auf neues Ziel
- Der Agent weiß nicht was seine Aktionen „bedeuten“
- Viele Zustände und Aktionen
- Lernen dauert mitunter sehr lange



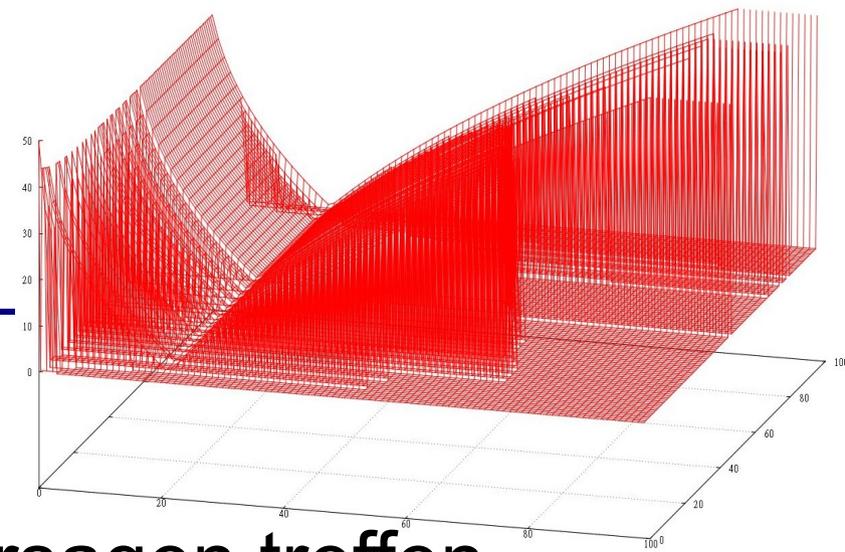
Seite 8

Die Idee



- Erlernen des Modells
- Der Agent hat eine „Vorstellung“ von dem was seine Aktionen „bewirken“
- Abhängigkeiten erkennen

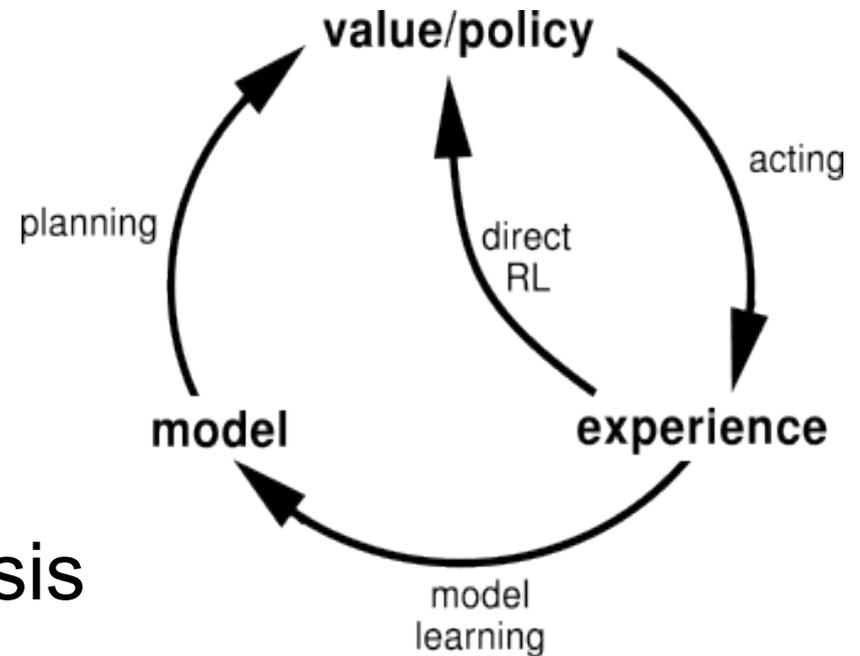
Die Idee



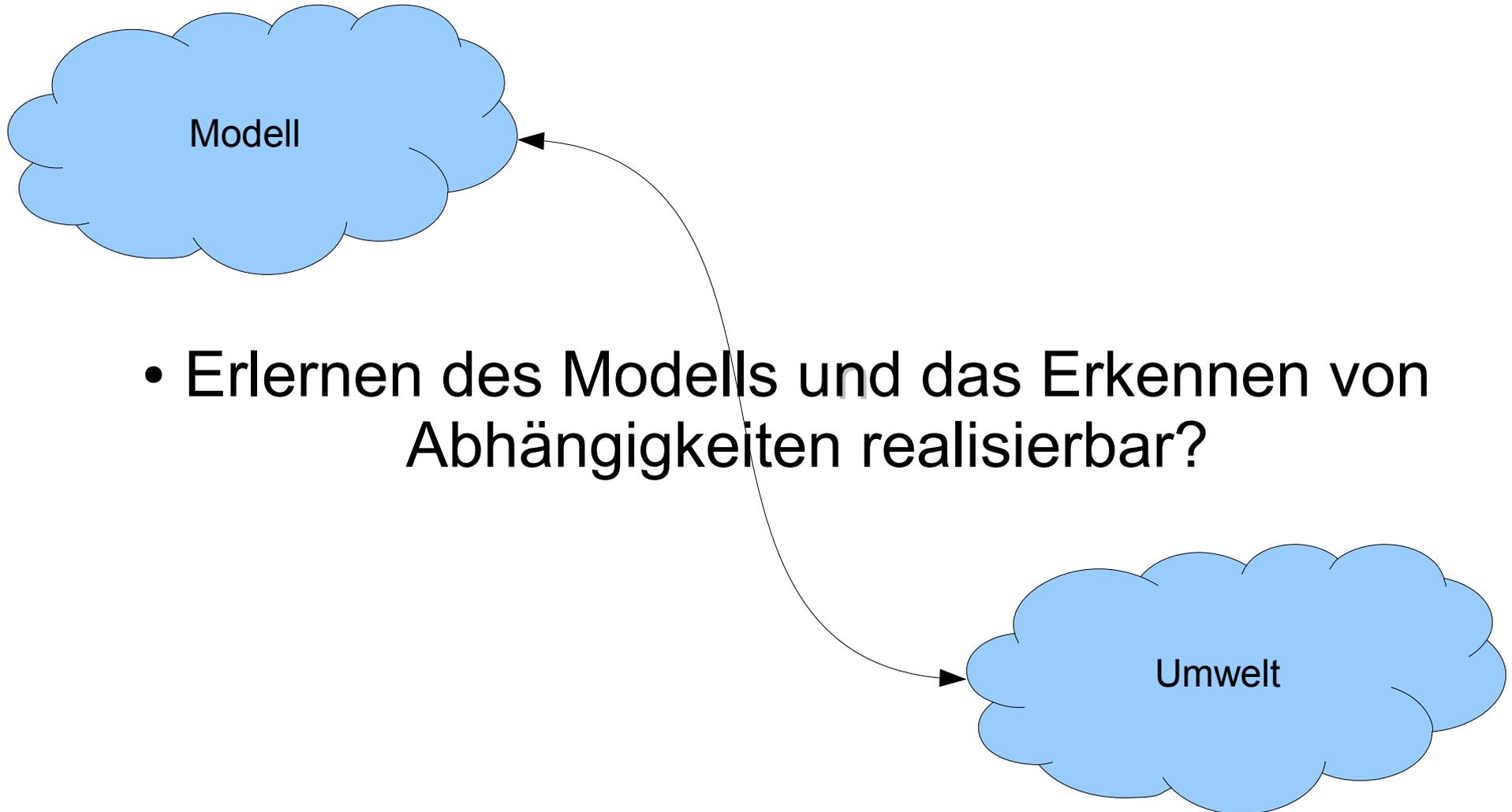
- Kennen des Modells → Vorhersagen treffen
- Stetiges optimieren des Modells
- Annahmen von bekannten Status auf unbekannte

Bestehender Ansatz

- Dyna Learning
- Lernen, planen, agieren
- Besuchte Zustände als Basis
- Eingabe: Zustand & Aktion
- Vorhersage: Folgezustand und dessen Belohnung



Das Risiko



- Erlernen des Modells und das Erkennen von Abhängigkeiten realisierbar?

Seite 12

Ausblick

- Konzeption eines neuen Lernverfahrens was die Vorteile von Reinforcement Learning und modellbasiertem Lernen verbindet
- Dyna analysieren und eventuell als Basis

Ausblick

- Simulation des Lernens mit geeigneter Simulationsumgebung (z.B. TORCS)
- Übertragung der Lernergebnisse auf FAUST-Fahrzeuge
- Nutzung der Lernalgorithmen in anderen Kontexten

Vielen Dank für Ihre Aufmerksamkeit



Literatur

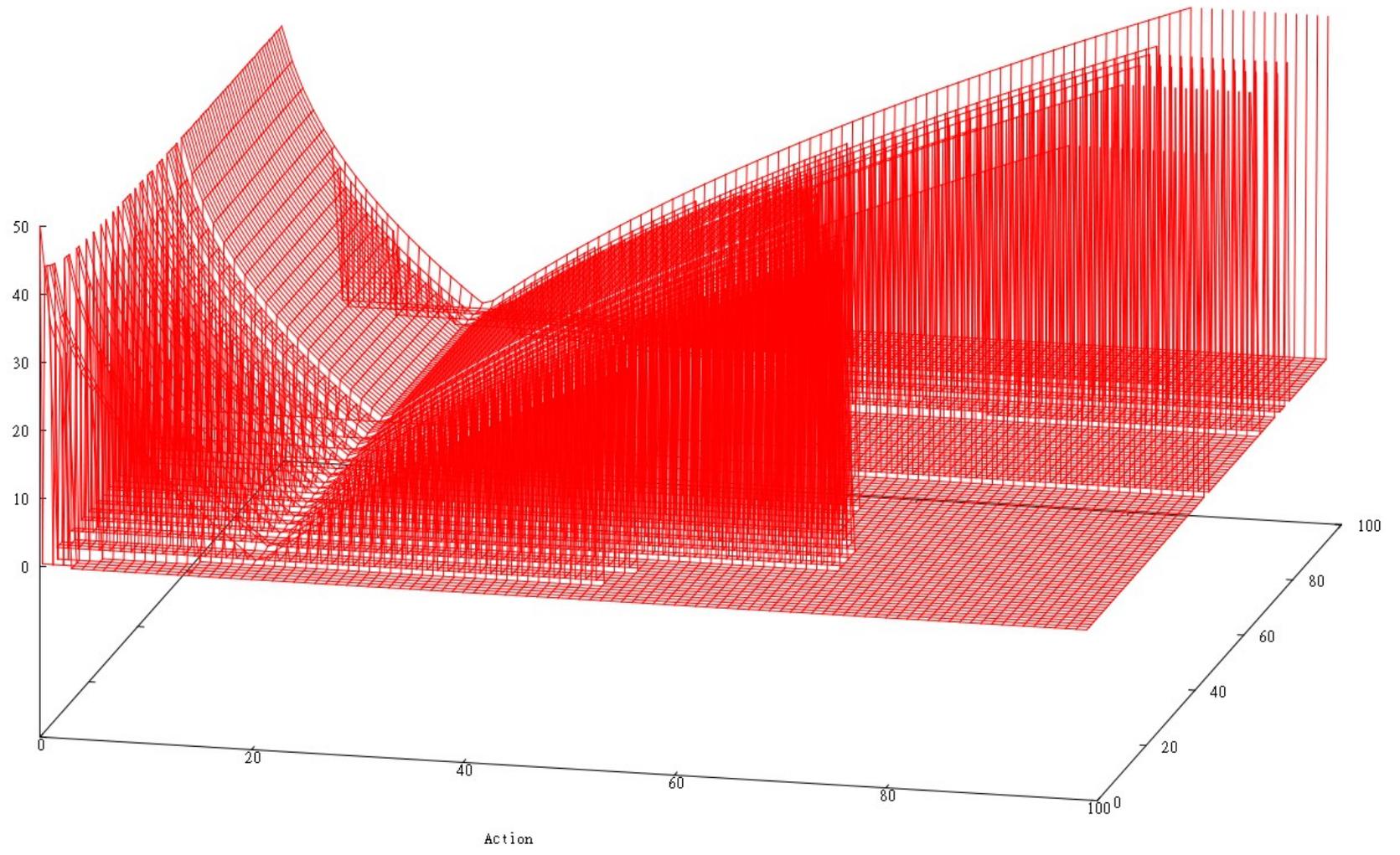
- Technik und Technologie 2 – Lernen, Prof. Dr. Michael Neitzke, 2010
- Reinforcement Learning: A Tutorial, Mance E. Harmon & Stephanie S. Harmon, 1996
- Maschinelles Lernen, Ethem Alpaydin, 2008
- Reinforcement Learning eine kurze Einführung, Tobias Jung, 2003
- Dyna, an Integrated Architecture for Learning, Planning, and Reacting, Richard S. Sutton, 1991
- Reinforcement Learning: An Introduction, Richard S. Sutton & Andrew G. Barto, 2005
<http://webdocs.cs.ualberta.ca/~sutton/book/ebook/the-book.html>
- Infoseite zu FAUST, <http://www.informatik.haw-hamburg.de/faust.html>, Stand 23.11.2010
- Q-Learning by Examples, Kardi Teknomo,
<http://people.revoledu.com/kardi/tutorial/ReinforcementLearning/index.html>, Stand 25.11.2010

Grafiken

- http://tevami.com/gallery/honda_asimo.jpg
- <http://webdocs.cs.ualberta.ca/~sutton/book/ebook/figtmp63.png>
- <http://blog.rpi-virtuell.net/gallery/2/klasse1-300.jpg>
- <http://skitterbot.com/images/cm-5.jpg>
- <http://www.dedoimedo.com/images/games/games-torcs-3.jpg>
- <http://ucgmikebennett.files.wordpress.com/2009/05/computer-question-mark-21107393.jpg>
- <http://i.techrepublic.com.com/blogs/ones-and-zeros.jpg>

"glearner10.csv" matrix

Reward



State