

CARIAD Planer – AI Driving Style

1 | Motivation und Ziele

- Erzeugung eines natürlichen, erwartungskonformen Fahrstils
- Anpassung des Fahrstils durch Tunen der Kostenfunktion eines suchbasierten Planers

Teilziele

1. Automatisierung des Tuning Prozesses durch das Nutzen von Fahrdemonstrationen
2. Erlernen eines dynamischen Situationsverständnisses
3. Dynamische und situative Anpassung des Fahrstils während Realfahrten

2 | Methode



Trajektoriengeneration

- Suchbasierter Planer erzeugt für das Ego Fahrzeug tausende kinematisch umsetzbare Trajektorien (im obigen Bild farbige dargestellt).
- Für alle Trajektorien werden Kostenmerkmale berechnet, die beispielsweise die Ego Kinematik (Beschleunigungen, Ruck, Geschwindigkeit) und Distanzen zu Objekten und Infrastruktur beschreiben.

Auswahl optimaler Trajektorie

- Auswahl einer Trajektorie über gewichtete Summe aller Kostenmerkmale
- Die Gewichtung wird über die Methode *Inverse Reinforcement Learning* erlernt.
- Das Wunschverhalten wird durch Demonstrationen gezeigt.
- Der Automatisierte Lernprozess über (Deep) Inverse Reinforcement Learning ermöglicht die Generalisierung über eine Vielzahl von Situationen.
- Statt einer konstanten Kostenfunktion für alle Situationen werden vom Modell situationsabhängige Kostenfunktionsgewichte generiert.

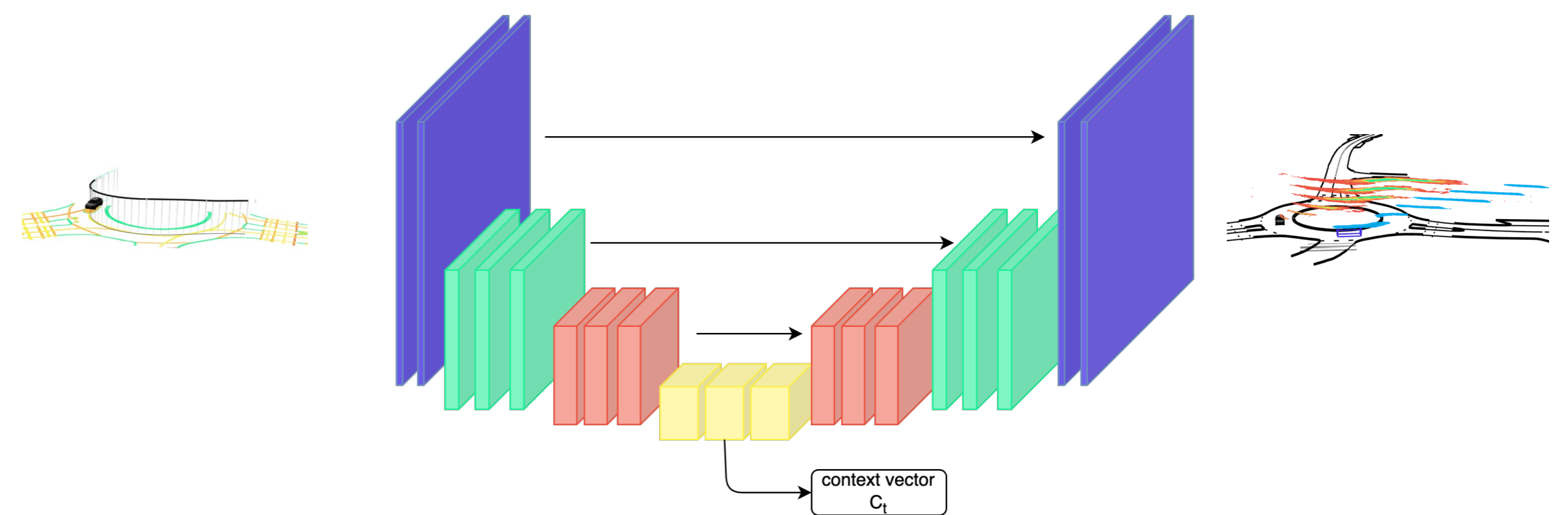
Situationsverständnis Modul

- Erzeugung eines natürlichen, erwartungskonformen Fahrstils kann nur durch gute Erkennung der aktuellen Situation erreicht werden.
- Entwicklung einer Deep Learning Methode, die den Situationskontext unter Berücksichtigung von anderen Verkehrsteilnehmern erlernen kann.
- Dieser Situationskontext wird genutzt, um die Parameter der Kostenfunktion situativ anzupassen.

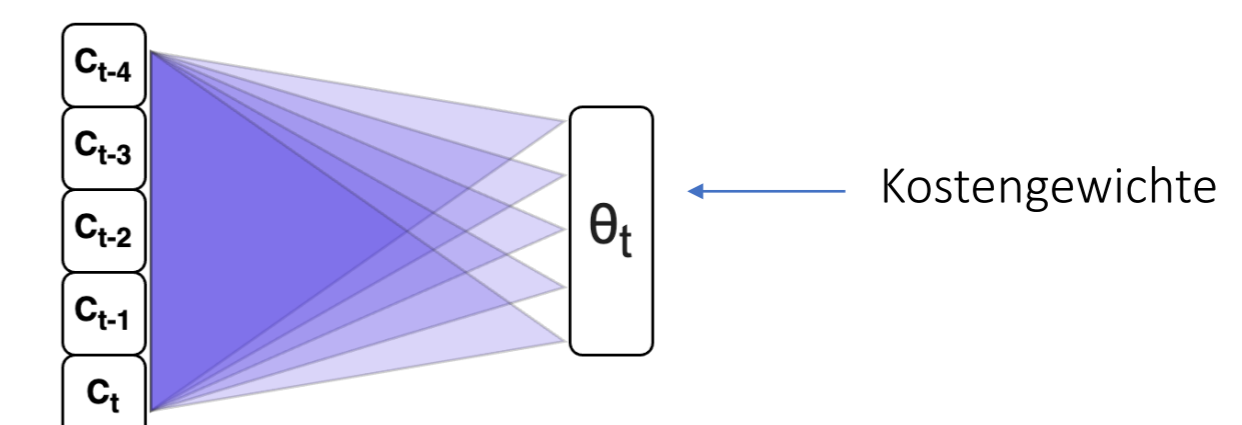
Situative Kostengewichte (Deep IRL)

- Transformer basiertes Neuronales Netz generiert ausgehend von aktueller Situation (kodiert im Kontext Vektor C) dynamisch Kostengewichte.
- Historie von Kontext Vektoren erlaubt zusätzliches Situationsverständnis.
- Training des Modells durch Inverse Reinforcement Learning: Kostenfunktion wird iterativ so angepasst, dass Demonstration möglichst wahrscheinlich ausgewählt werden würde während die Entropie möglichst groß bleibt.

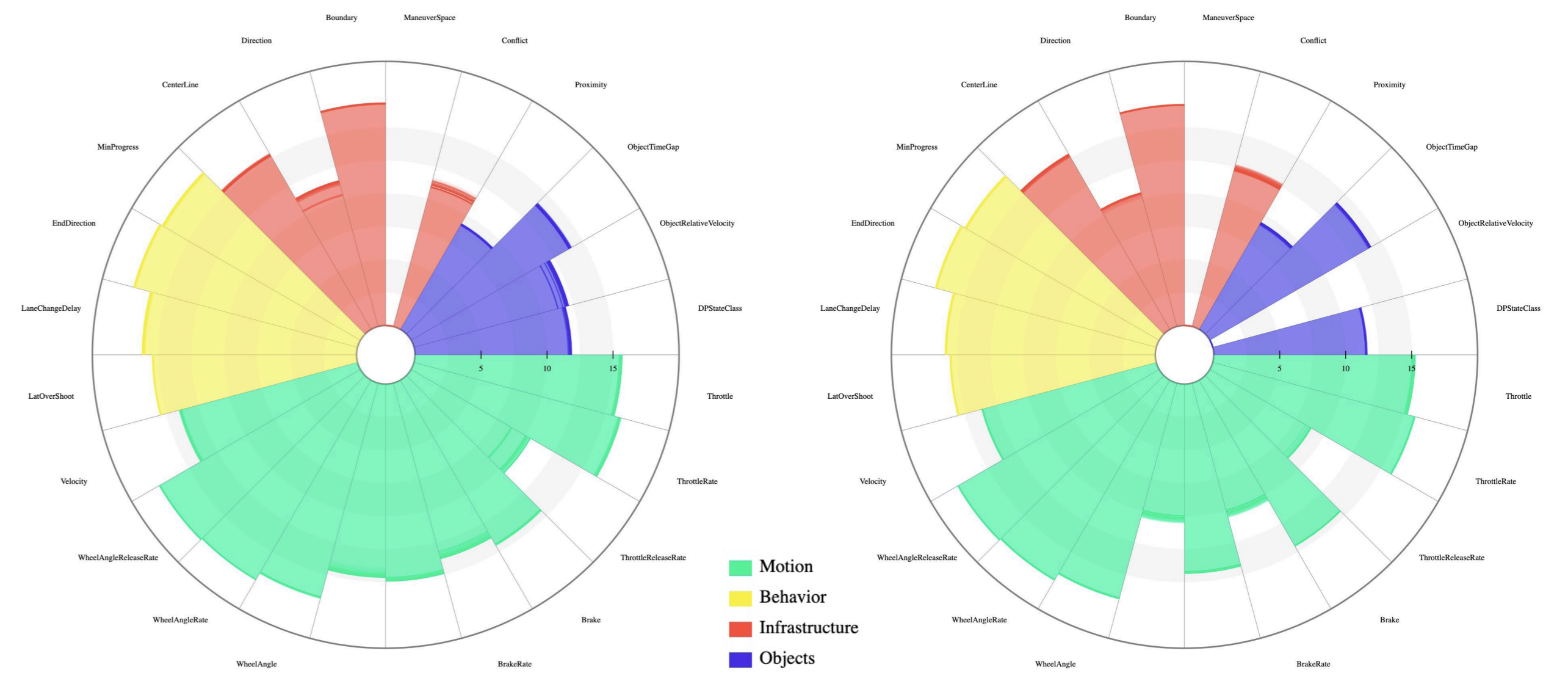
Schritt 1: Pretraining (Situationsverständnis)



Schritt 2: Inverse Reinforcement Learning (Training des Fahrstils)



3 | Ergebnisse



Je nach Situation ergeben sich unterschiedliche Kostengewichte (links und rechts im Vergleich). Dadurch wird der Fahrstil an die gegebene Situation (viele Objekte, Kreisverkehr, Autobahnfahrt) angepasst.

4 | Diskussion

- Die Methodik ermöglicht es dem hier dargestellten suchbasierenden Planer den Fahrstil anhand von Demonstrationen zu erlernen.
- Es wird eine Reduktion von situativen Verhaltensimplementierungen erreicht, welche die Generalisierungsfähigkeit einer Fahrfunktion einschränken.
- Die gelernten Kostenfunktionen wurden in urbanen Realfahrten der begleitenden Studien erfolgreich eingesetzt.

Literatur

- [1] Rosbach, S., James, V., Großjohann, S., Homoceanu, S., & Roth, S. (2019, November). Driving with style: Inverse reinforcement learning in general-purpose planning for automated driving. In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 2658-2665). IEEE.
- [2] Rosbach, S., James, V., Großjohann, S., Homoceanu, S., Li, X., & Roth, S. (2020, May). Driving style encoder: Situational reward adaptation for general-purpose planning in automated driving. In 2020 IEEE international conference on robotics and automation (ICRA) (pp. 6419-6425). IEEE.
- [3] Rosbach, S., Leupold, S. M., Großjohann, S., & Roth, S. (2023). Pixel State Value Network for Combined Prediction and Planning in Interactive Environments. arXiv preprint arXiv:2310.07706.