13th International Conference on Applied Informatics
Eszterházy Károly Catholic University
Eger, Hungary, February 19–21, 2026

# Modeling Dynamic Reaction Time in Car-Following via Transformer-based Inverse Reinforcement Learning

**László Mészáros**[a], **András Hajdu**[a]

[a]University of Debrecen, Faculty of Informatics
meszaros.laszlo@inf.unideb.hu
hajdu.andras@inf.unideb.hu

## Abstract

Accurate modeling of human reaction time is an important component of autonomous driving systems, particularly in mixed traffic environments, where discrepancies between human and autonomous vehicle decision-making can result in unsafe, inefficient, or uncomfortable interactions.

Traditional car-following models often rely on fixed reaction time parameters to describe longitudinal vehicle dynamics. However, human driving behavior is inherently adaptive, as cognitive delay varies based on environmental factors and vehicle states. Unlike prior Transformer-based Inverse Reinforcement Learning (IRL) approaches that implicitly assume fixed temporal relevance, we explicitly model human reaction time as a dynamic, state-dependent temporal attention distribution learned from data [2]. This study proposes a novel approach using Transformer-based Deep IRL to capture and interpret the dynamic nature of human reaction time, building upon existing interpretable planning architectures [3]. By leveraging the attention mechanism of Transformers, we provide a transparent method to visualize how a driver's temporal focus shifts under different traffic conditions.

We utilize the NGSIM US-101 dataset [4], extracting trajectory pairs of leading and following vehicles. The proposed architecture consists of a Transformer Encoder to capture temporal dependencies and a Generative adversarial imitation learning (GAIL) framework [1]. In our formulation, reaction time is not treated as a fixed delay parameter but emerges as a latent, interpretable variable derived from the temporal attention weights of the Transformer. This interpretability layer identifies which past time steps are most influential for the current acceleration de-

cision, allowing the model to recover the underlying driver intent.

Analysis of attention maps reveals a bimodal distribution of cognitive focus, typically peaking at approximately 0.7 seconds and 0.2 seconds into the past. Our primary finding is the dynamic shift of the focus peak relative to the speed of the vehicle. Regression analysis demonstrates a correlation: as vehicle speed increases, the temporal focus shifts closer to the present. This suggests that at higher speeds, human-like agents exhibit shorter reaction delays, allocating greater cognitive weight to more recent stimuli to mitigate increased safety risks. This behavior provides a more authentic representation of human "alertness" than static models, aligning with the goals of safe and interpretable human-like planning.

This research demonstrates that Transformer-based IRL models can provide explainable insights into the underlying cognitive processes. This enables autonomous driving policies that adapt their temporal decision-making horizon in a human-consistent manner, improving both safety and behavioral predictability in mixed traffic. Our findings facilitate the development of robust autonomous systems that are both predictable to human drivers and safe in complex environments.

# References

[1] J. Ho, S. Ermon: *Generative adversarial imitation learning*, Advances in neural information processing systems 29 (2016).

[2] J. Nan, W. Deng, R. Zhang, R. Zhao, Y. Wang, J. Ding: *Car-Following Behavior Modeling With Maximum Entropy Deep Inverse Reinforcement Learning*, IEEE Transactions on Intelligent Vehicles 9.2 (2024), pp. 3998–4010, DOI: 10.1109/TIV.2023.3335218.

[3] J. Nan, R. Zhang, G. Yin, W. Zhuang, Y. Zhang, W. Deng: *Safe and Interpretable Human-Like Planning With Transformer-Based Deep Inverse Reinforcement Learning for Autonomous Driving*, IEEE Transactions on Automation Science and Engineering 22 (2025), pp. 12134–12146, DOI: 10.1109/TASE.2025.3539340.

[4] U.S. Federal Highway Administration: *Next Generation Simulation (NGSIM) Vehicle Trajectories and Supporting Data*, https://catalog.data.gov/dataset/next-generation-simulation-ngsim-vehicle-trajectories-and-supporting-data, Accessed: 2026-01-23, 2016.