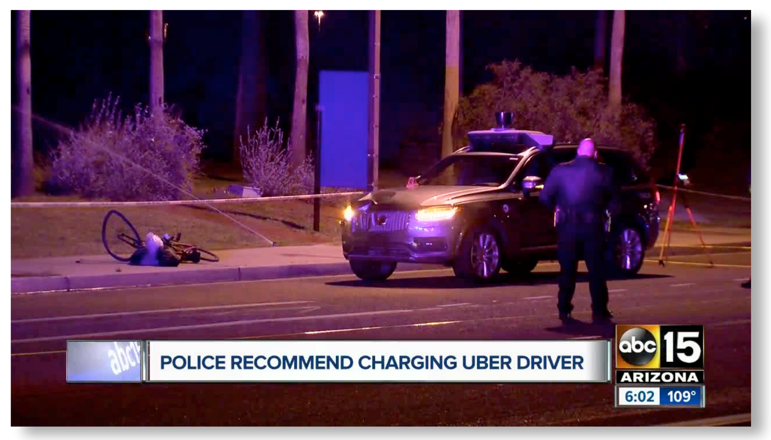
🤔 Think Deep and Fast ⚡ | Learning Neural NOD from Inverse Dynamic Games for Split-Second Interactions

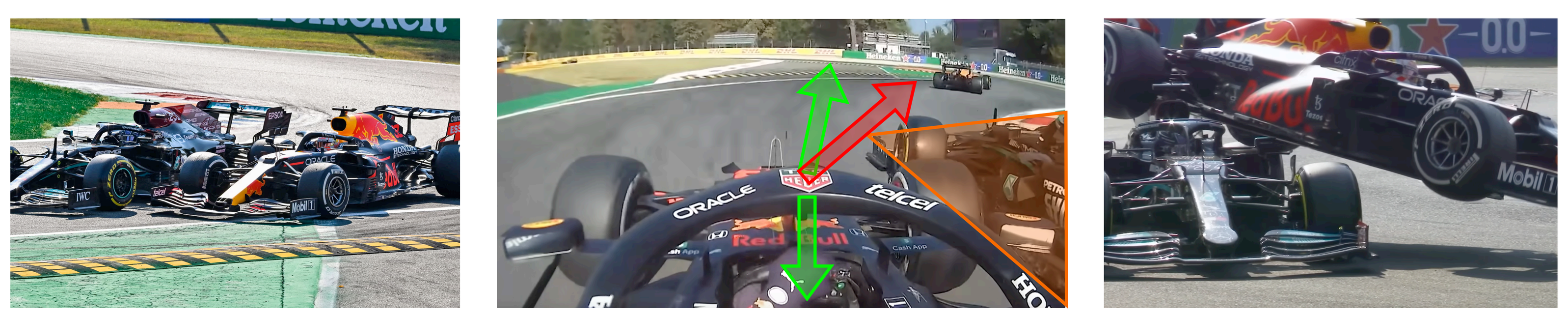
Haimin Hu, Jaime Fisac, Naomi Leonard, Deepak Gopinath, Jonathan DeCastro, Guy Rosman

SAFE ROBOTICS LABORATORY · Leonard LABORATORY · PRINCETON UNIVERSITY · TOYOTA RESEARCH INSTITUTE

Motivation: Time-Sensitive Interactions



How can robots learn to make safe and performant decisions in time-sensitive strategic interactions?



Incident involving world champions L. Hamilton and M. Verstappen during the 2021 Formula 1 Italian Grand Prix

Key idea: Planning game-theoretic robot motion with guarantees of indecision breaking.

Background: Opinion-Driven Games

Nonlinear Opinion Dynamics (NOD)

Let $z^i \in \mathbb{R}^{N_{\theta^i}}$ be an opinion vector that represents agent i's preference over N_{θ^i} discrete options

$z^i_\ell > 0$ (< 0) means that agent i favors (disfavors) option ℓ

$z^i_\ell = 0$ means that agent i is neutral toward option ℓ

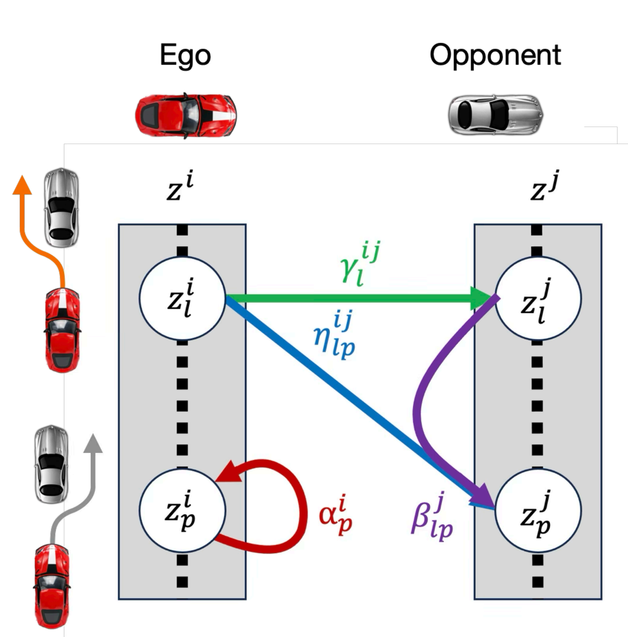
$$\dot{z}^i = -d^i z^i + b^i + \lambda^i S^i_z(z^i)$$
$$\dot{\lambda}^i = -m\lambda^i + S^i_\lambda(\mathbf{z}^i)$$

where

self-reinforcement · inter-agent, same option

$$S^{i,\ell}_z(z^i) = S_1\left(\alpha^i_\ell z^i_\ell + \sum_{j\in\mathcal{I}_a\setminus\{i\}} \gamma^{ij}_\ell z^j_\ell\right) + \sum_{p\in\mathcal{I}_{\theta_i}\setminus\{\ell\}} S_2\left(\beta^i_{\ell p} z^i_p + \sum_{j\in\mathcal{I}_a\setminus\{i\}} \eta^{ij}_{\ell p} z^j_p\right)$$

intra-agent coupling · inter-agent, different option



NOD properties:

- Guaranteed opinion formation (i.e., indecision breaking)
- Exponential rate of opinion formation
- Efficient and scalable computation

⚠️ Not trivial to construct one

Opinion-Guided Dynamic Games (ODG)

$$\dot{z} = g_z(z, x, u),\ z \in \mathbb{R}^{n_z} \rightarrow \phi^i \leftarrow \text{softplus}(z^i)$$

Ego cost

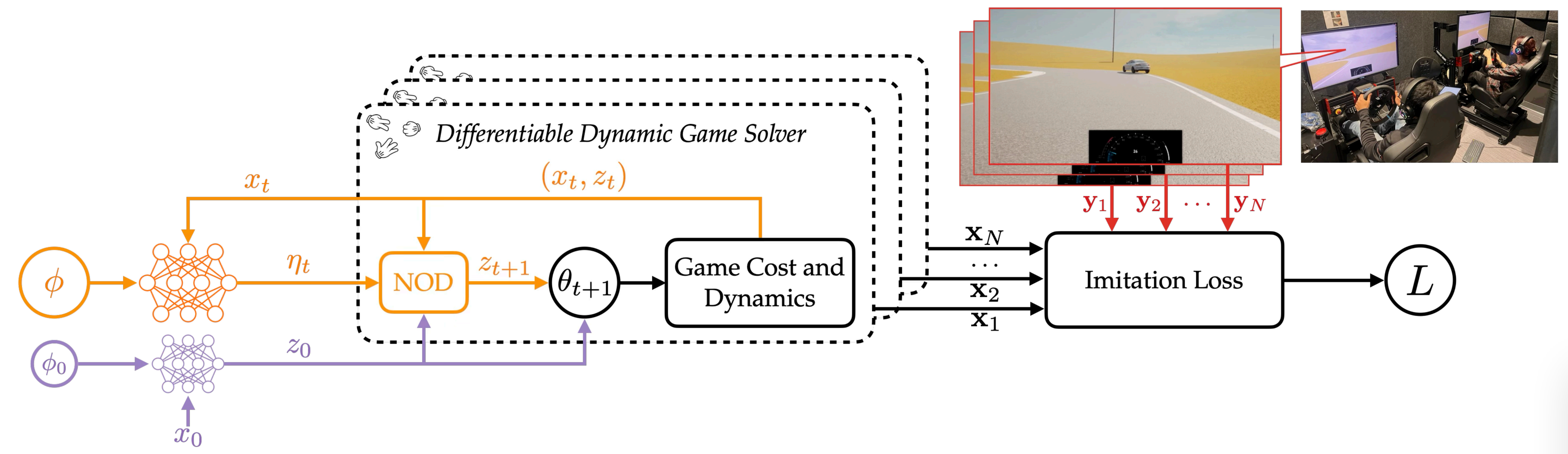
$$c^e(x_t, u_t; \phi^e) = \boxed{\phi^e_{\text{stay}}}\, c^e_{\text{stay}}(x_t, u_t) + \boxed{\phi^e_{\text{overtake}}}\, c^e_{\text{overtake}}(x_t, u_t) + c^e_{\text{other}}(x_t, u_t)$$

$$c^a(x_t, u_t; \phi^a) = \boxed{\phi^a_{\text{block}}}\, c^a_{\text{block}}(x_t, u_t) + c^a_{\text{other}}(x_t, u_t)$$

Rival cost *prediction*

- Opinion value adaptively changes game payoff coefficients
- Closes the loop between opinion evolution and physical interaction

Neural NOD for Interactive Robotics



Neural NOD Example

$$\dot{z}^e_1 = -d^e_1(x,u)z^e_1 + \lambda^e(x,u)\left[S\left(\alpha^e_1(x,u)z^e_1\right) + S\left(\beta^e_{12}(x,u)z^e_2 + \delta^{ea}_{11}(x,u)z^a_1\right)\right]$$
$$\dot{z}^e_2 = -d^e_2(x,u)z^e_2 + \lambda^e(x,u)\left[S\left(\alpha^e_2(x,u)z^e_2\right) + S\left(\beta^e_{21}(x,u)z^e_1 + \delta^{ea}_{21}(x,u)z^a_1\right)\right]$$
$$\dot{z}^a_1 = -d^a_1(x,u)z^a_1 + \lambda^a(x,u)\left[S\left(\alpha^a_1(x,u)z^a_1\right) + S\left(\sum_{i=1,2}\delta^{ae}_{1i}(x,u)z^e_i\right)\right]$$

$$(x,u) \longrightarrow \text{[DNN]} \longrightarrow \alpha(\cdot), \beta(\cdot), \delta(\cdot), \lambda(\cdot), d(\cdot)$$

Inverse Game Training

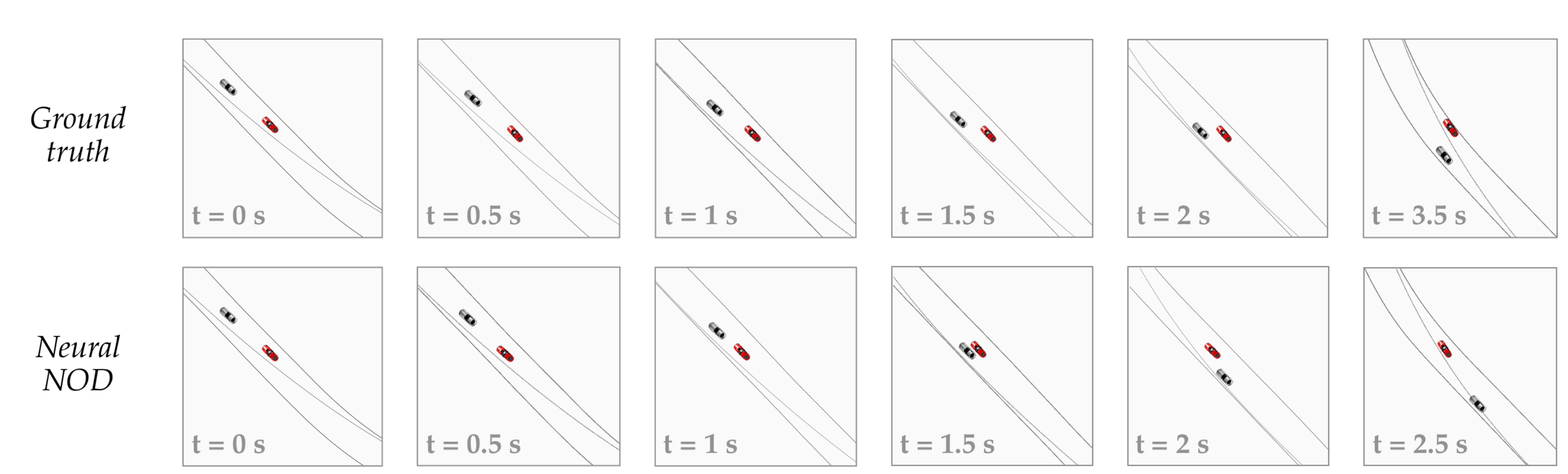
$$\max_{\phi, \phi_0, \{(\mathbf{x}_n, \mathbf{u}_n)\}_{n\in[N]}} \quad L := \sum_{n=1}^N p(\mathbf{y}_n \mid \mathbf{x}_n, \mathbf{u}_n)$$
$$\text{s.t.} \quad (\mathbf{x}_n, \mathbf{u}_n) \in \Gamma(\phi, \phi_0), \quad \forall n \in [N]$$

Approach & Features:

- **Game-aware synthesis:** Learn a Neural NOD from expert demonstrations using inverse dynamic game training
- **Situational awareness:** The learned DNN outputs NOD parameters that change adaptively based on the observation

Theoretical Properties:

- With non-zero bias, the neutral opinion is not an equilibrium; the model is ultrasensitive at $z = 0$ (pitchfork unfolds)
- With zero bias, it is possible to analytically construct α and λ such that the neutral opinion is an unstable equilibrium

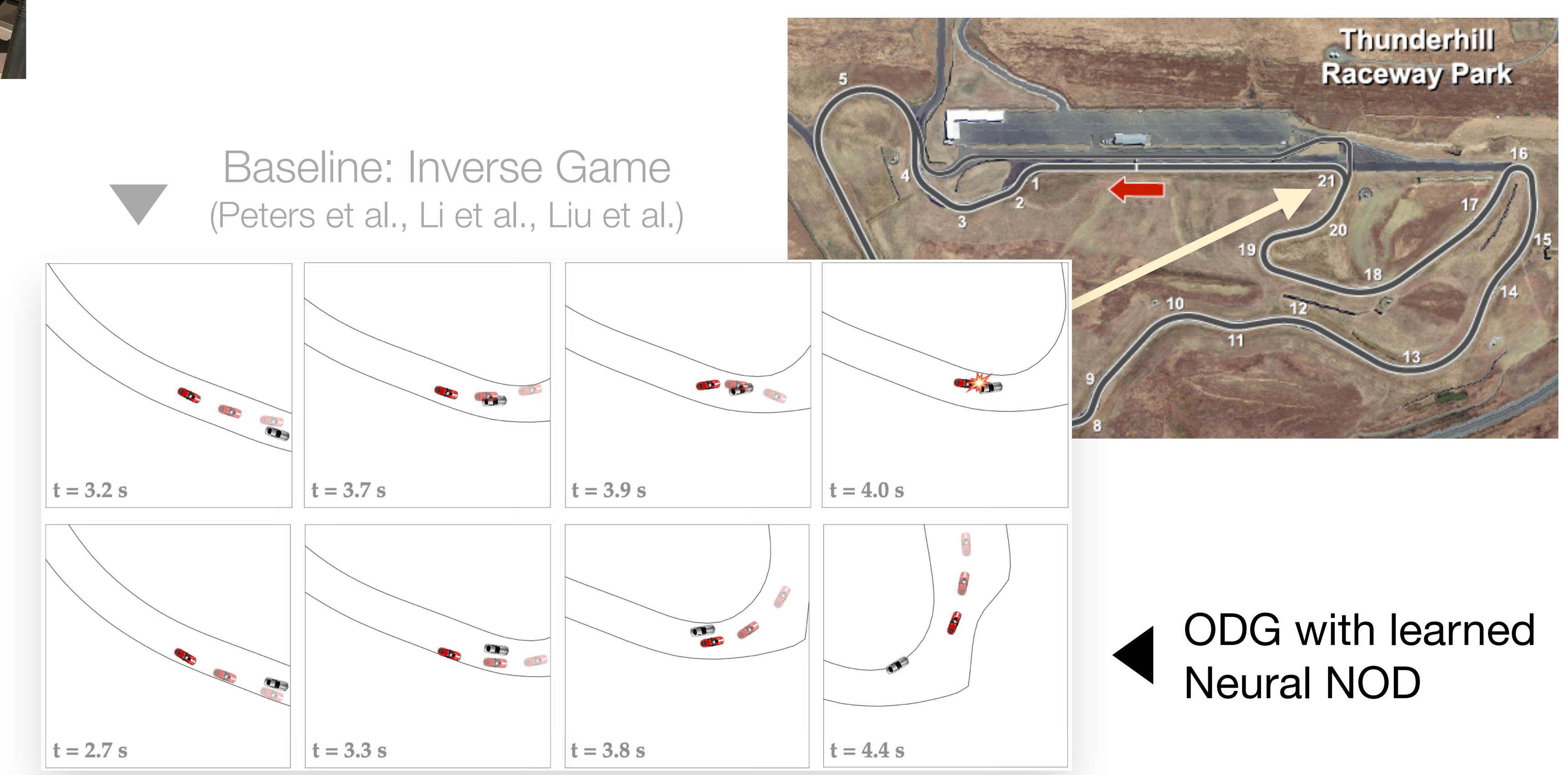


IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION · ICRA ATLANTA 2025

👀 *Related Papers:* KL Games · ODG

Case Studies: High-Speed Car Racing

Qualitative Results: Overtaking at Thunderhill Raceway T21



Quantitative Results: Duel and Endurance Races

Results from 100 Duel Races:

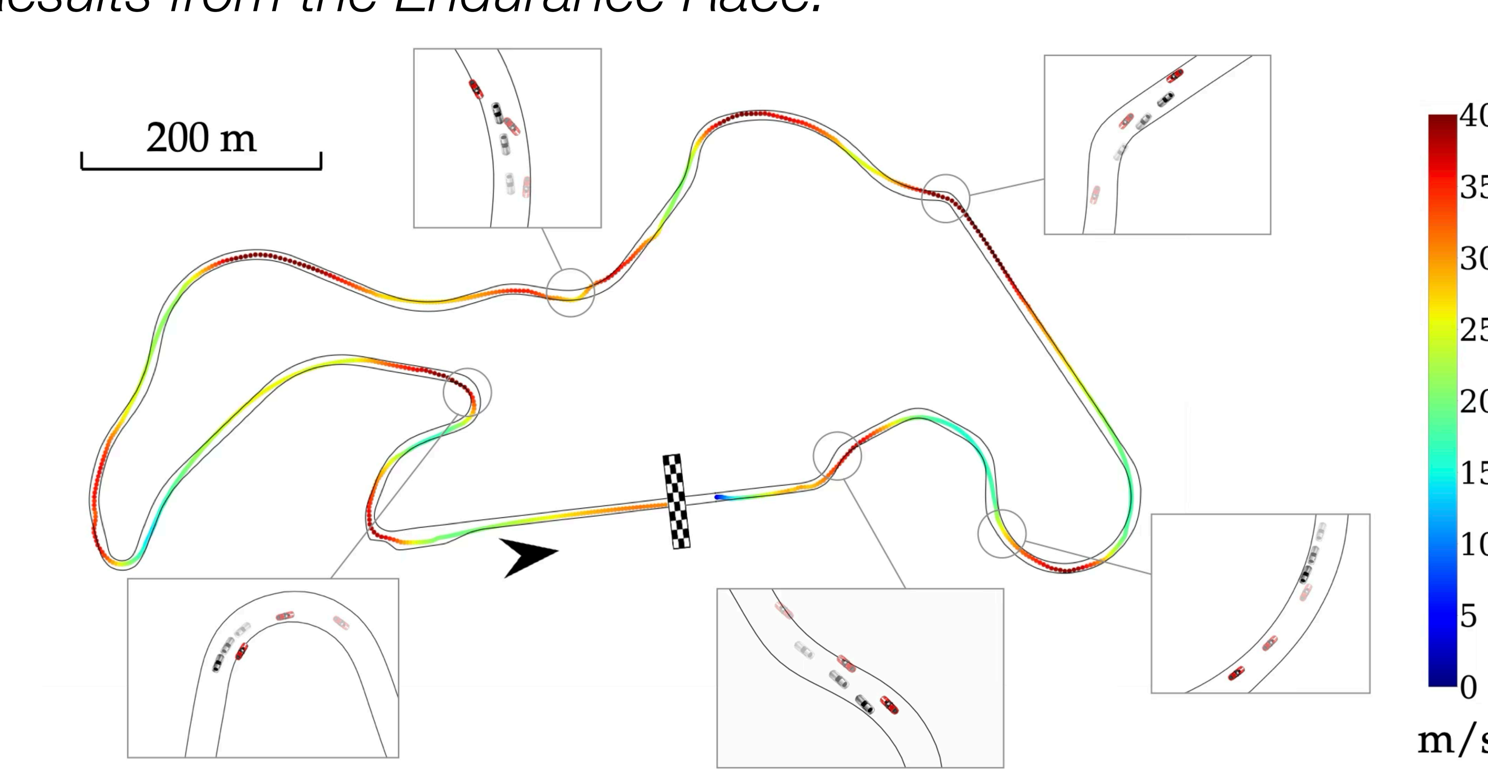
Synthetic Data

Metrics/Methods	Neural NOD	MLP	BC
Safe rate	**91%** (▲9%)	82%	66%
Overtaking rate	**76.92%** (▲14.72%)	62.20%	60.61%
Avg. leading distance	**7.60 m** (▲68.89%)	4.50 m	3.64 m

Human Data

Metrics/Methods	Neural NOD	MLP	BC
Safe rate	**81%** (▲3%)	78%	62%
Overtaking rate	**82.72%** (▲28.87%)	53.85%	75.81%
Avg. leading distance	**15.94 m** (▲10.01%)	14.49 m	12.35 m

Results from the Endurance Race:



Metrics/Methods	Neural NOD	MLP	BC
Lap Time	**101.4 s** (▼11.21%)	114.2 s	110.1 s
Collisions	**0**	2	3