Learning to Play Trajectory Games Against Opponents with Unknown Objectives

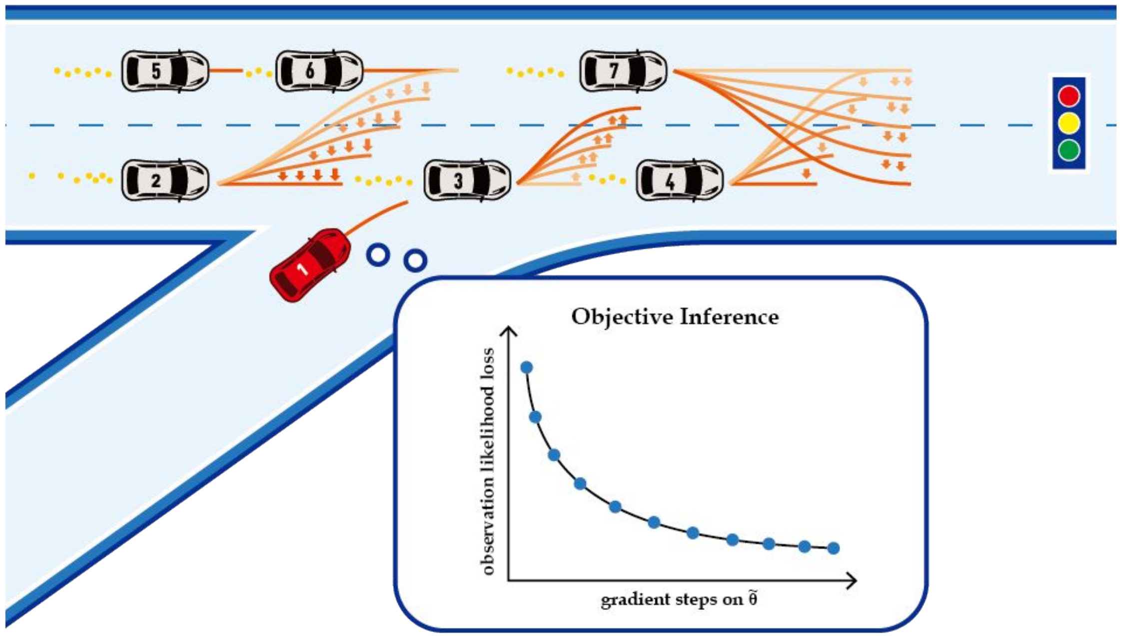
Xinjie Liu

Cognitive Robotics (CoR)

x.liu-47@student.tudelft.nl

TU Delft

Motivation



- Autonomous driving: robots need to reason about interactions
- Our perspective: dynamic games (**explicit** modeling of interactions, **simultaneous** predicting and planning)
- Alternative: predict-then-plan

"Forward" Dynamic Games

An N-player open-loop Nash game as coupled trajectory optimization:

$$\forall i \in [N] \begin{cases} \min\limits_{X^i, U^i} & J^i(\mathbf{X}, U^i; \theta^i) \\ \text{s.t.} & x^i_{t+1} = f^i(x^i_t, u^i_t), \forall t \in [T-1] \\ & x^i_1 = \hat{x}^i_1 \\ & {}^p g^i(X^i, U^i) \geq 0 \\ & {}^s g(\mathbf{X}, \mathbf{U}) \geq 0 \end{cases}$$

cost function

system dynamics

initial states

private inequalities

shared inequalities

Solution: generalized Nash equilibrium (GNE)

$$J^i(\mathbf{X}^*, U^{i*}; \theta^i) \leq J^i((X^i, \mathbf{X}^{\neg i*}), U^i; \theta^i)$$

No unilateral change of controls can reduce a player's costs.

TU Delft

"Forward" Dynamic Games

An N-player open-loop Nash game as coupled trajectory optimization:

$$\forall i \in [N] \begin{cases} \min_{X^i, U^i} & J^i(\mathbf{X}, U^i; \theta^i) \\ \text{s.t.} & x^i_{t+1} = f^i(x^i_t, u^i_t), \forall t \in [T-1] \\ & x^i_1 = \hat{x}^i_1 \\ & {}^p g^i(X^i, U^i) \geq 0 \\ & {}^s g(\mathbf{X}, \mathbf{U}) \geq 0 \end{cases}$$

cost function

system dynamics

initial states

private inequalities

shared inequalities

Solution: generalized Nash equilibrium (GNE)

$$J^i(\mathbf{X}^*, U^{i*}; \theta^i) \leq J^i((X^i, \mathbf{X}^{\neg i*}), U^i; \theta^i)$$

No unilateral change of controls can reduce a player's costs.

TU Delft

"Forward" Dynamic Games

An N-player open-loop Nash game as coupled trajectory optimization:

Partially observable stochastic game (POSG), generally **intractable**!!

$$\forall i \in [N] \begin{cases} \min\limits_{X^i, U^i} & J^i(\mathbf{X}, U^i; \theta^i) \\ \text{s.t.} & x_{t+1}^i = f^i(x_t^i, u_t^i), \forall t \in [T-1] \\ & x_1^i = \hat{x}_1^i \\ & {}^p g^i(X^i, U^i) \geq 0 \\ & {}^s g(\mathbf{X}, \mathbf{U}) \geq 0 \end{cases}$$

cost function

system dynamics

initial states

private inequalities

shared inequalities

Solution: generalized Nash equilibrium (GNE)

$$J^i(\mathbf{X}^*, U^{i*}; \theta^i) \leq J^i((X^i, \mathbf{X}^{\neg i*}), U^i; \theta^i)$$

No unilateral change of controls can reduce a player's costs.

TU Delft

Inverse Games

$$\begin{aligned}
\max_{\theta,\mathbf{X},\mathbf{U}} \quad & p(\overbrace{\mathbf{Y}}^{\text{observation}} \mid \mathbf{X}, \mathbf{U}) \\
\text{s.t.} \quad & (\mathbf{X}, \mathbf{U}) \text{ is a GNE of the game } \Gamma(\theta)
\end{aligned}$$

Applications (w.r.t. explicit modeling of the interactions):

- Online interaction with other agents (POSG approximation)
- Trajectory prediction
- Tuning of the ego-agent's controller to match desired behavior (similar to inverse RL)

Inverse Games

$$\max_{\theta,\mathbf{X},\mathbf{U}} \quad p(\overbrace{\mathbf{Y}}^{\text{observation}} \mid \mathbf{X}, \mathbf{U})$$

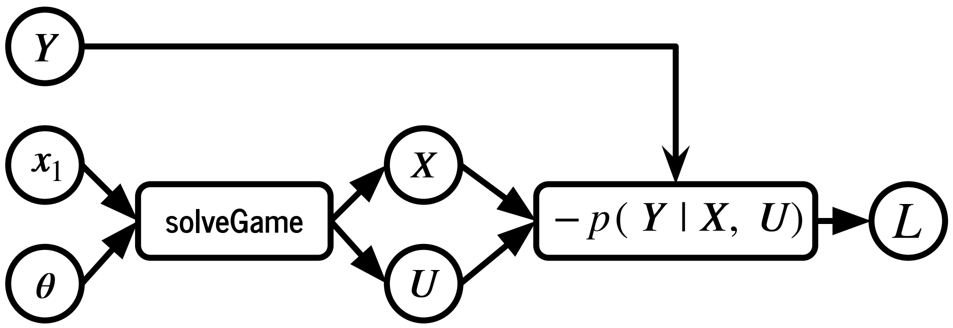
$$\text{s.t.} \quad \underbrace{\boxed{(\mathbf{X}, \mathbf{U}) \text{ is a GNE of the game } \Gamma(\theta)}}_{\text{optimality conditions of a forward game}}$$

Challenge: how to efficiently encode the equilibrium constraints?

- Highly nonlinear
- Naive encoding violates constraint qualification ($\lambda^\top g(\mathbf{X}, \mathbf{U}) = 0$)
- Real-time computation

Approach

The Forward Computation Graph



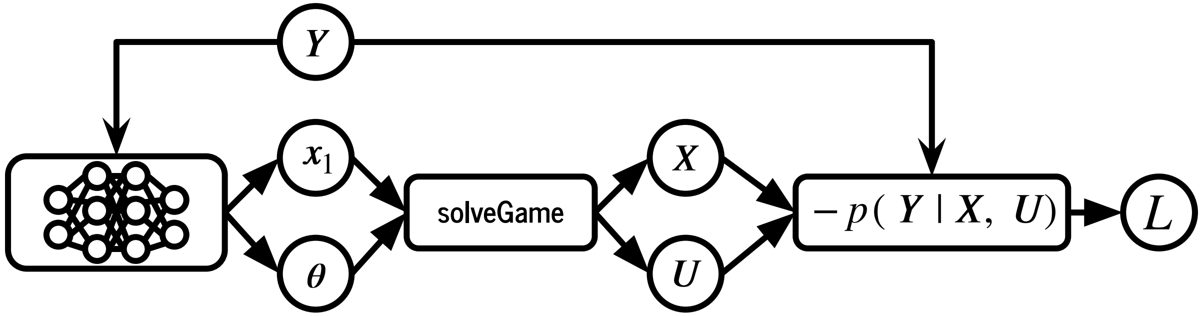
This entire computation graph can be made differentiable!

⇒ We can update estimates of θ and x1 via gradient descent on the loss function.

TU Delft

Approach

Differentiable Games | Extensions



The gradient signal can be back-propagated to a neural network, which learns to predict the game parameters.

Approach: "Forward" Games as Mixed Complementarity Problems (MCPs)

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$
\begin{aligned}
z_j^* = \ell_j, F_j(z^*) \geq 0 \\
\ell_j < z_j^* < u_j, F_j(z^*) = 0 \\
z_j^* = u_j, F_j(z^*) \leq 0.
\end{aligned}
$$

Solver reference: S. P. Dirkse and M. C. Ferris, "The PATH solver: A nommonotone stabilization scheme for mixed complementarity problems," 1995.

Approach: "Forward" Games as Mixed Complementarity Problems (MCPs)

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$\begin{aligned} z_j^* = \ell_j, F_j(z^*) \geq 0 \\ \ell_j < z_j^* < u_j, F_j(z^*) = 0 \\ z_j^* = u_j, F_j(z^*) \leq 0. \end{aligned}$$

Optimality conditions of a game can be cast as the solution to an equivalent MCP!

$$\forall i \in [N] \begin{cases} \nabla_{(X^i, U^i)} \mathcal{L}^i(\mathbf{X}, \mathbf{U}, \mu^i, {}^p\lambda^i, {}^s\lambda; \theta) = 0 \\ 0 \leq {}^p g^i(X^i, U^i) \perp {}^p\lambda^i \geq 0 \end{cases}$$
$$h(\mathbf{X}, \mathbf{U}; \hat{\mathbf{x}}_1) = 0$$
$$0 \leq {}^s g(\mathbf{X}, \mathbf{U}) \perp {}^s\lambda \geq 0,$$

$$\longrightarrow$$

$$z = \begin{bmatrix} X^1 \\ U^1 \\ \vdots \\ X^N \\ U^N \\ \boldsymbol{\mu} \\ {}^p\lambda^1 \\ \vdots \\ {}^p\lambda^N \\ {}^s\lambda \end{bmatrix} \quad F(z;\theta) = \begin{bmatrix} \nabla_{(X^1, U^1)} \mathcal{L}^i(\mathbf{X}, \mathbf{U}, \mu^i, {}^p\lambda^1, {}^s\lambda; \theta) \\ \vdots \\ \nabla_{(X^N, U^N)} \mathcal{L}^N(\mathbf{X}, \mathbf{U}, \mu^i, {}^p\lambda^N, {}^s\lambda; \theta) \\ h(\mathbf{X}, \mathbf{U}; \hat{\mathbf{x}}_1) \\ {}^p g^1(X^1, U^1) \\ \vdots \\ {}^p g^N(X^N, U^N) \\ {}^s g(\mathbf{X}, \mathbf{U}) \end{bmatrix} \quad \ell = \begin{bmatrix} -\infty \\ \vdots \\ -\infty \\ -\infty \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad u = \begin{bmatrix} \infty \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \infty \end{bmatrix}$$

TU Delft

Solver reference: S. P. Dirkse and M. C. Ferris, "The PATH solver: A nommonotone stabilization scheme for mixed complementarity problems," 1995.

Approach: Differentiation Through Mixed Complementarity Problems (MCPs)

$\theta \rightarrow$ ***SolveMCP*** $\rightarrow z^*$

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$
\begin{aligned}
z_j^*(\theta) = \ell_j, F_j(z^*(\theta); \theta) \geq 0 \\
\ell_j < z_j^*(\theta) < u_j, F_j(z^*(\theta); \theta) = 0 \\
z_j^*(\theta) = u_j, F_j(z^*(\theta); \theta) \leq 0.
\end{aligned}
$$

TU Delft

Approach: Differentiation Through Mixed Complementarity Problems (MCPs)

$\theta \longrightarrow \boxed{\textit{SolveMCP}} \longrightarrow z^*$

Assumption: **strong complementarity**

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$\begin{aligned} z_j^*(\theta) = \ell_j, F_j(z^*(\theta); \theta) &\geq 0 \\ \ell_j < z_j^*(\theta) < u_j, F_j(z^*(\theta); \theta) &= 0 \\ z_j^*(\theta) = u_j, F_j(z^*(\theta); \theta) &\leq 0. \end{aligned}$$

TU Delft

Approach: Differentiation Through Mixed Complementarity Problems (MCPs)

$\theta \longrightarrow \boxed{\textbf{\textit{SolveMCP}}} \longrightarrow z^*$

Assumption: **strong complementarity**

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$
\begin{aligned}
z_j^*(\theta) = \ell_j, F_j(z^*(\theta); \theta) > 0 \\
\ell_j < z_j^*(\theta) < u_j, F_j(z^*(\theta); \theta) = 0 \\
z_j^*(\theta) = u_j, F_j(z^*(\theta); \theta) < 0.
\end{aligned}
$$

Approach: Differentiation Through Mixed Complementarity Problems (MCPs)

$\theta \longrightarrow$ ***SolveMCP*** $\longrightarrow z^*$

Assumption: **strong complementarity**

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$\left.\begin{aligned} z_j^*(\theta) = \ell_j, F_j(z^*(\theta);\theta) > 0 \\ \ell_j < z_j^*(\theta) < u_j, F_j(z^*(\theta);\theta) = 0 \\ z_j^*(\theta) = u_j, F_j(z^*(\theta);\theta) < 0. \end{aligned}\right\} \nabla_\theta \tilde{z}^* = 0$$

Approach: Differentiation Through Mixed Complementarity Problems (MCPs)

$\theta \longrightarrow \boxed{\textbf{\textit{SolveMCP}}} \longrightarrow z^*$

Assumption: **strong complementarity**

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$
\begin{aligned}
z_j^*(\theta) = \ell_j, F_j(z^*(\theta);\theta) > 0 \\
\boxed{\ell_j < z_j^*(\theta) < u_j, F_j(z^*(\theta);\theta) = 0} \\
z_j^*(\theta) = u_j, F_j(z^*(\theta);\theta) < 0.
\end{aligned}
\qquad \nabla_\theta \tilde{z}^* = 0
$$

Approach: Differentiation Through Mixed Complementarity Problems (MCPs)

$\theta \longrightarrow \boxed{\textit{SolveMCP}} \longrightarrow z^*$

Assumption: **strong complementarity**

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

$$z_j^*(\theta) = \ell_j, F_j(z^*(\theta);\theta) > 0$$

$$\boxed{\ell_j < z_j^*(\theta) < u_j, F_j(z^*(\theta);\theta) = 0}$$

$$z_j^*(\theta) = u_j, F_j(z^*(\theta);\theta) < 0.$$

$$\nabla_\theta \tilde{z}^* = 0$$

Implicit function theorem:

$$0 = \nabla_\theta \left[\bar{F}(z^*(\theta), \theta)\right] =$$

$$\nabla_\theta \bar{F} + (\nabla_{\bar{z}^*} \bar{F})(\nabla_\theta \bar{z}^*) + (\nabla_{\tilde{z}^*} \bar{F}) \underbrace{(\nabla_\theta \tilde{z}^*)}_{\equiv 0} \longrightarrow \nabla_\theta \bar{z}^* = -\left(\nabla_{\bar{z}^*} \bar{F}\right)^{-1} (\nabla_\theta \bar{F})$$

TU Delft

Approach: Differentiation Through Mixed Complementarity Problems (MCPs)

$\theta \longrightarrow \boxed{\textbf{SolveMCP}} \longrightarrow z^*$

Weak complementarity: subgradient
Invertibility: least-square solution

In a *Mixed Complementarity Problem*, we have decision variable $z \in \mathbb{R}^n$ and problem data $F(z) : \mathbb{R}^n \mapsto \mathbb{R}^n, \ell_j \in \mathbb{R} \cup \{-\infty\}, u_j \in \mathbb{R} \cup \{\infty\}, j \in [n]$. At the solution z^*, *one* of the following cases holds:

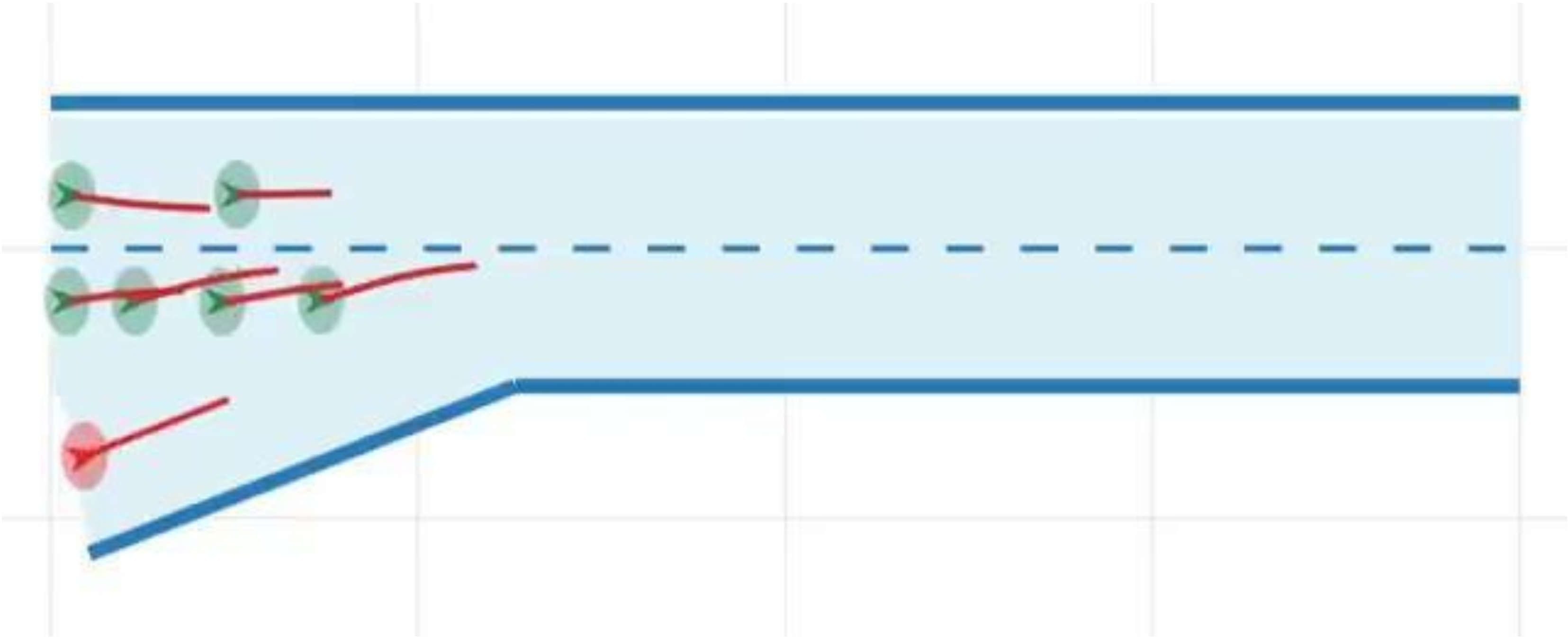
$$z_j^*(\theta) = \ell_j, F_j(z^*(\theta); \theta) > 0$$
$$\boxed{\ell_j < z_j^*(\theta) < u_j, F_j(z^*(\theta); \theta) = 0}$$
$$z_j^*(\theta) = u_j, F_j(z^*(\theta); \theta) < 0.$$

$$\nabla_\theta \tilde{z}^* = 0$$

Implicit function theorem:

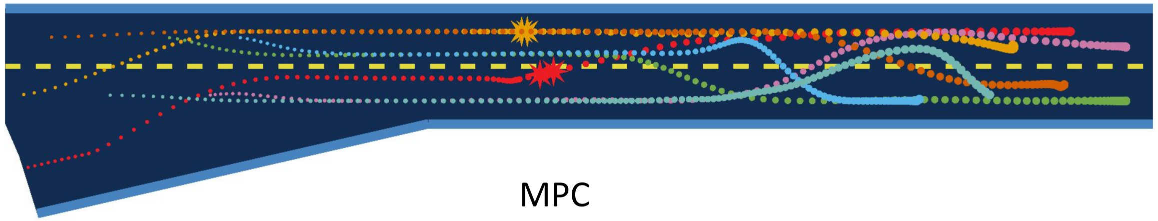
$$0 = \nabla_\theta \left[\bar{F}(z^*(\theta), \theta)\right] = \nabla_\theta \bar{F} + (\nabla_{\bar{z}^*} \bar{F})(\nabla_\theta \bar{z}^*) + (\nabla_{\tilde{z}^*} \bar{F}) \underbrace{(\nabla_\theta \tilde{z}^*)}_{\equiv 0} \longrightarrow \nabla_\theta \bar{z}^* = -\left(\nabla_{\bar{z}^*} \bar{F}\right)^{-1} (\nabla_\theta \bar{F})$$

Example: 7-player Highway Driving

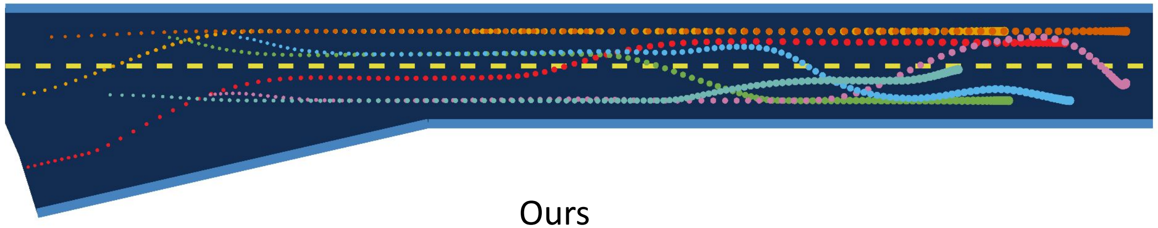


supplementary video: https://www.youtube.com/watch?v=f0KJuCC1Xyo

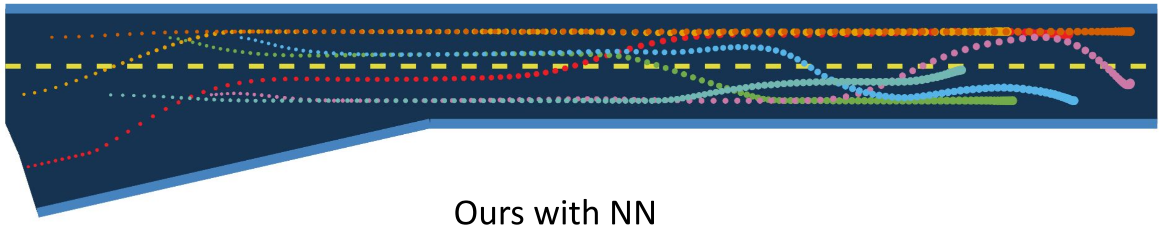
Example: 7-player Highway Driving



MPC



Ours



Ours with NN

TU Delft

Example: Guiding-Tracking Game on Jackals

2x

Example: Human-Robot Interaction

1x

TU Delft

Future Work

- Integrated end-to-end planning with perception module (picking up additional visual cues, such as gaze or body language, for inference)

- Robustness against uncertainty (reasoning about the inference confidence)

Collaborators

Lasse Peters

Javier Alonso-Mora

Thanks for your attention! :-D