On Game-Theoretic Planning with Unknown Opponents' Objectives

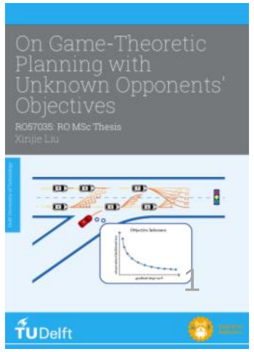
Xinjie Liu
MSc in Robotics

10 July 2023

Thesis committee:

Associate Professor Javier Alonso-Mora, Supervisor
Assistant Professor Laura Ferranti
Assistant Professor Luca Laurenti
Lasse Peters, Daily supervisor

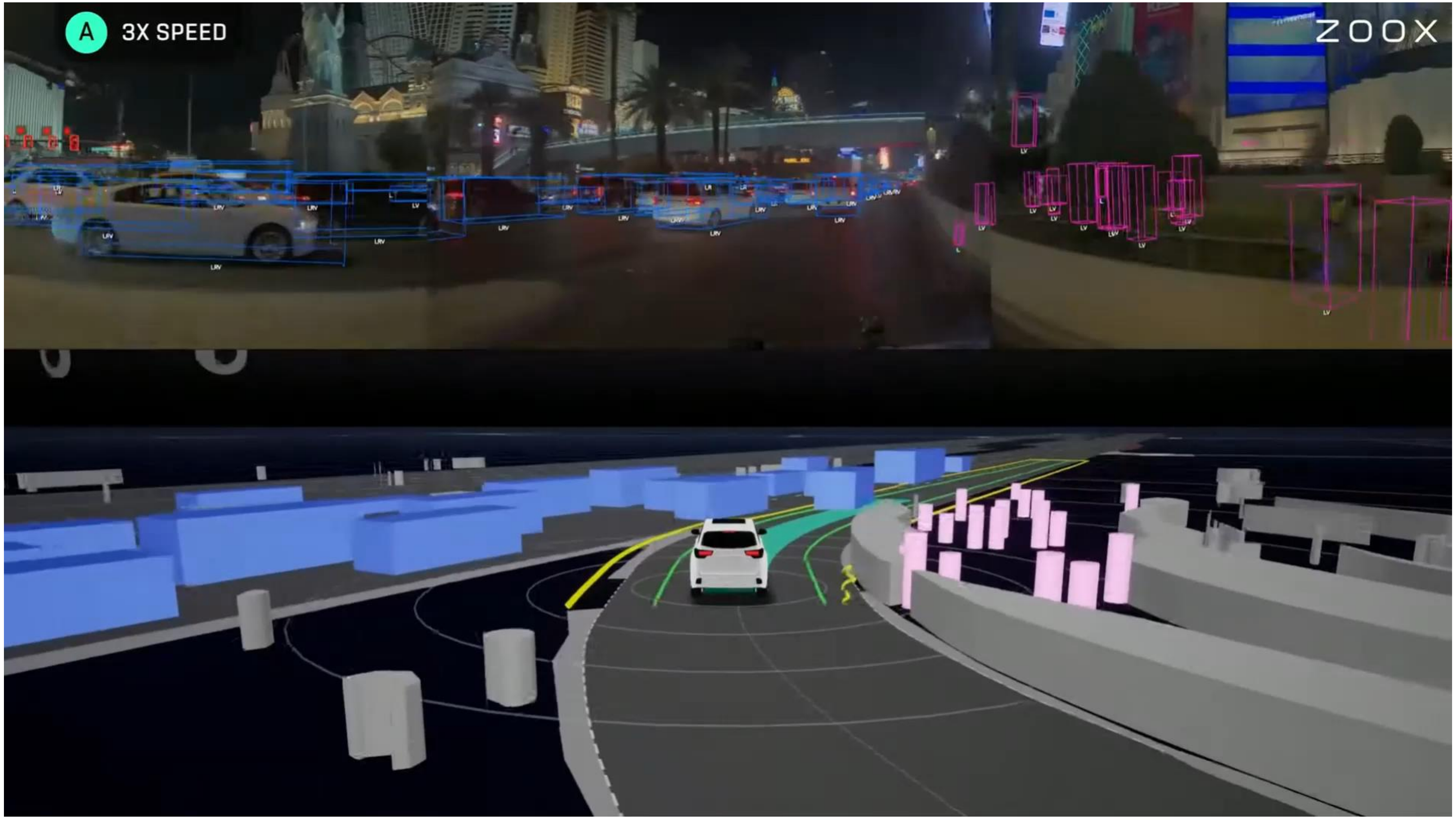
TU Delft | Cognitive Robotics | AUTONOMOUS MULTI-ROBOTS LAB Delft University of Technology



Multi-Agent Interaction

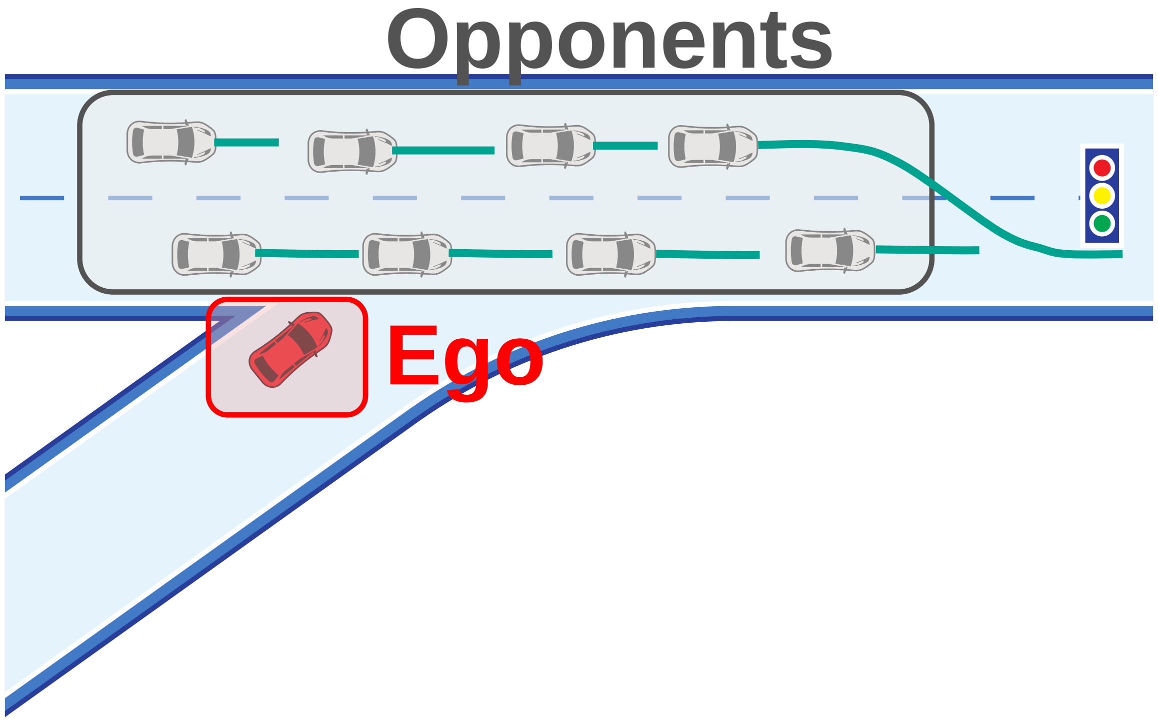
Video: Thomas Schlijper, 2018.

Multi-Agent Interaction



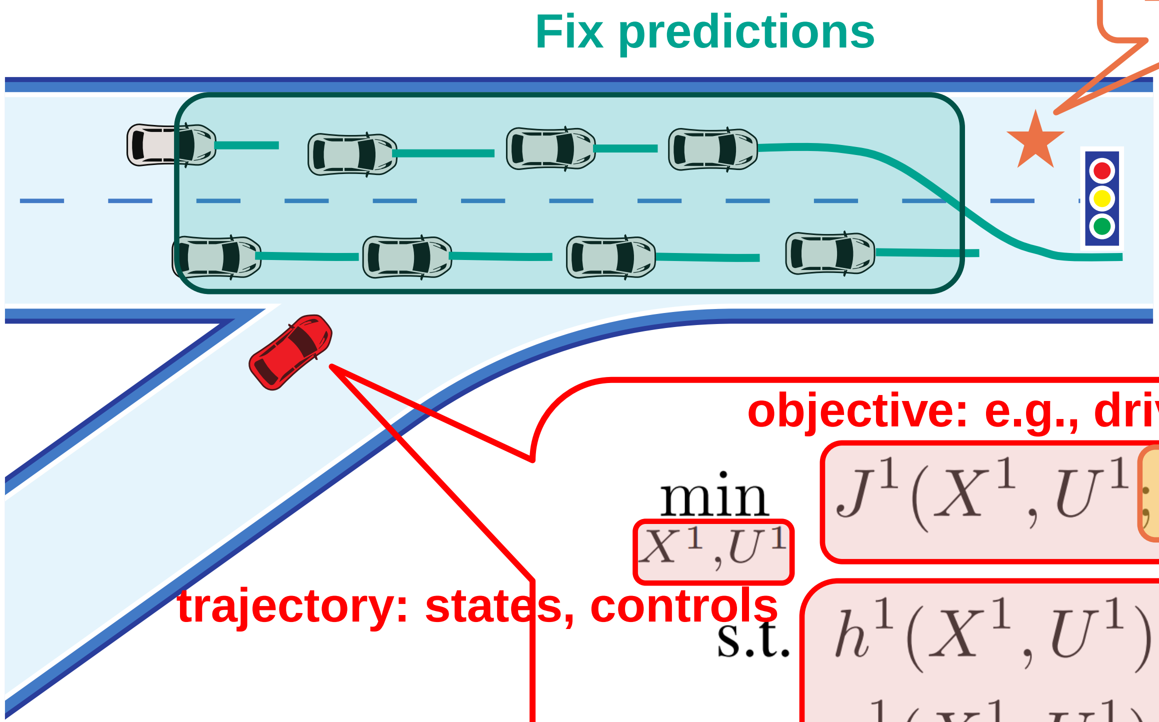
Video: Zoox, 2021.

Motivation



- Single-agent **optimal control**
  - A popular framework: “predict-then-plan”

Motivation



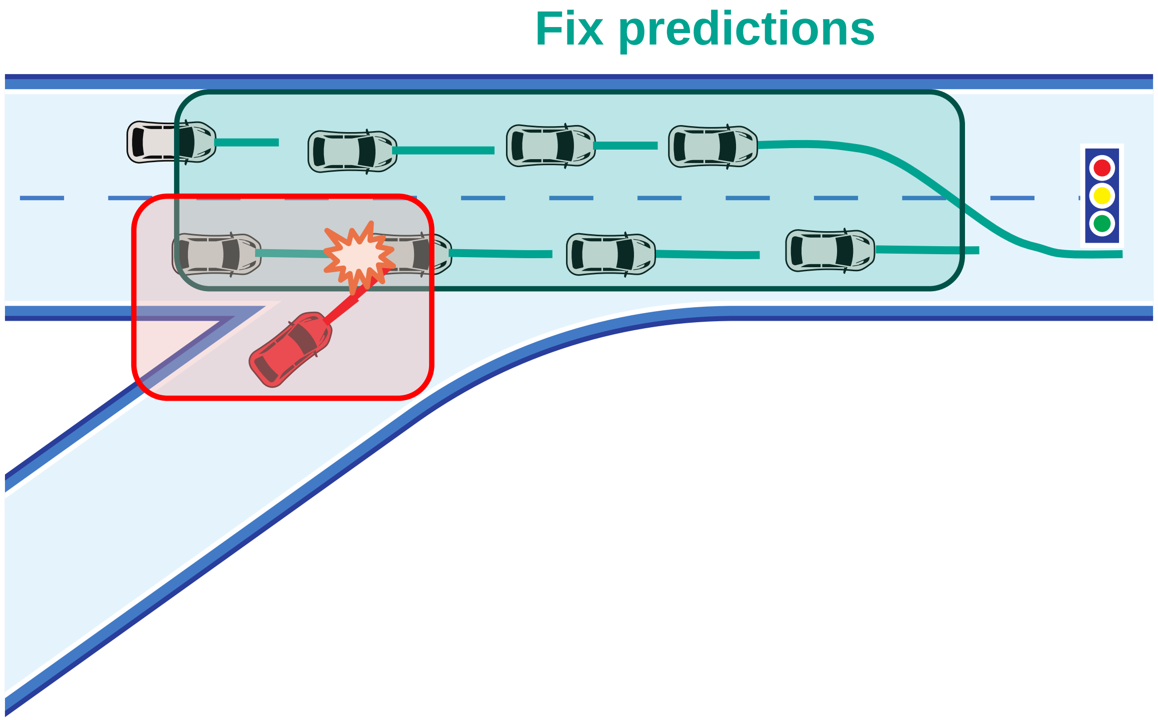
- Single-agent **optimal control**
  - A popular framework: “predict-then-plan”

$$\min_{X^1,U^1} J^1(X^1,U^1;\theta^1)$$
$$\text{s.t. } h^1(X^1,U^1)=0$$
$$g^1(X^1,U^1)\geq 0$$

e.g., desired speed, target lane

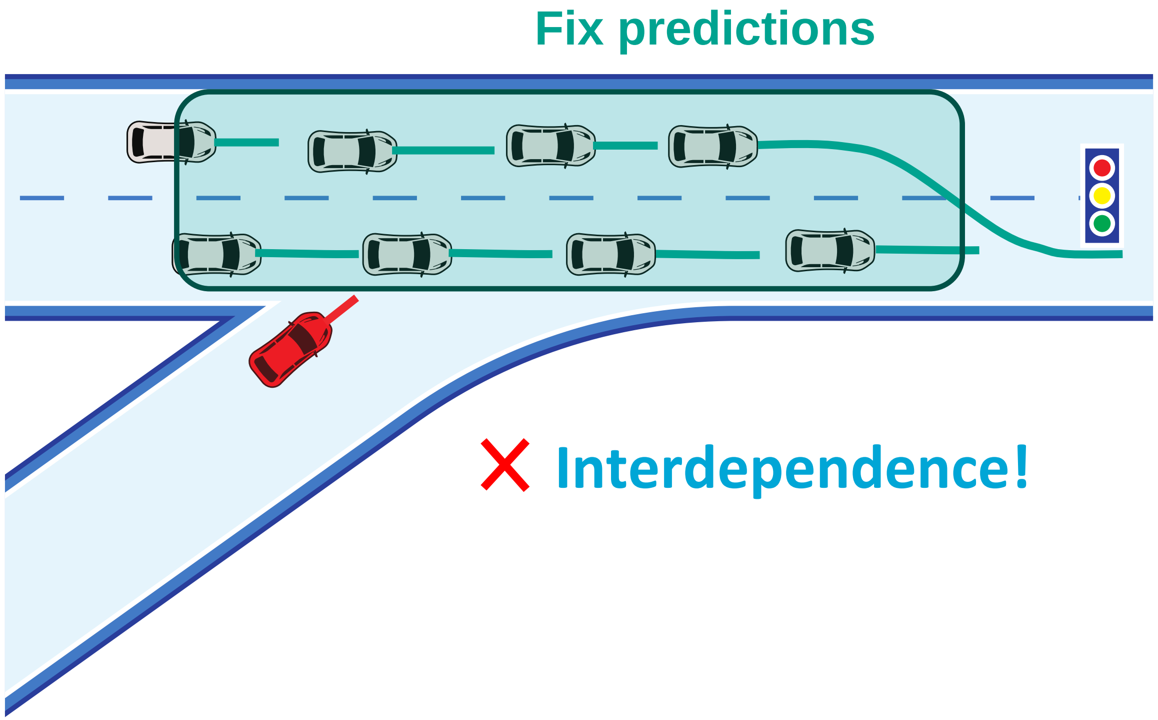
constraints, e.g., collision avoidance

Motivation



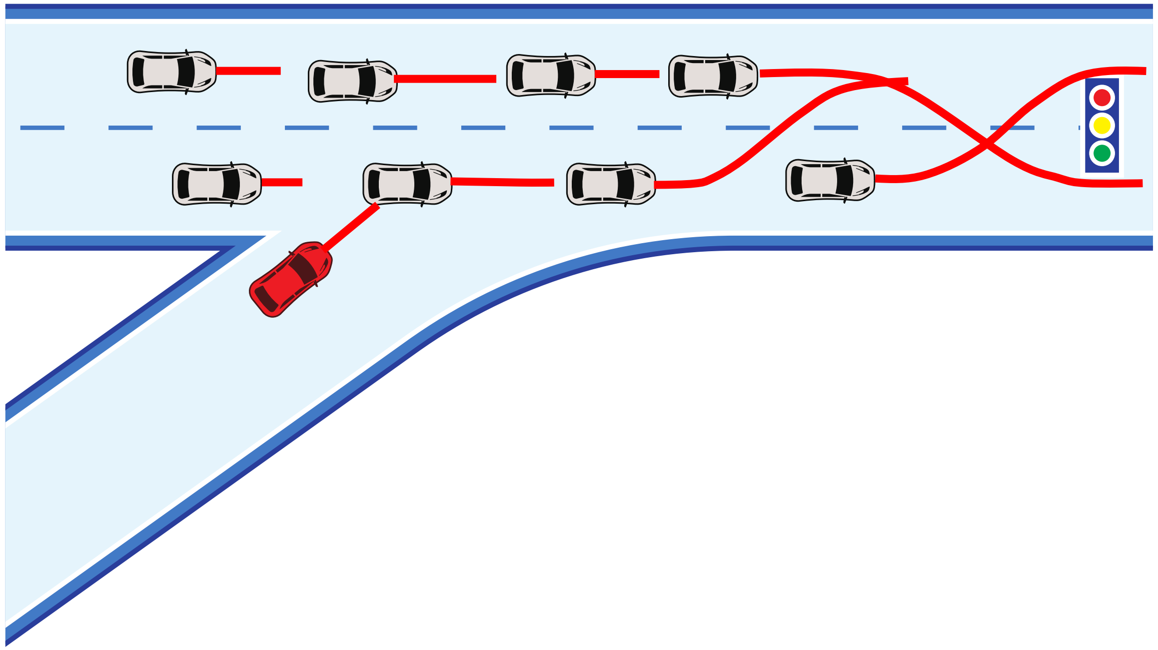
- Single-agent **optimal control**
  - A popular framework: “predict-then-plan”

Motivation



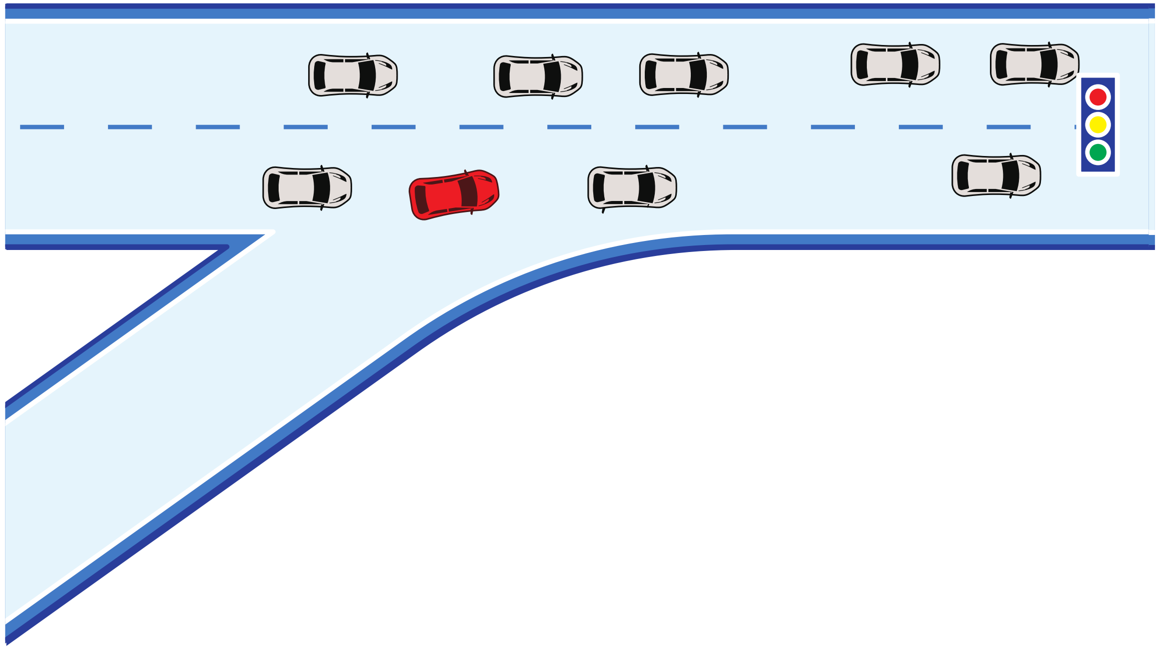
- Single-agent **optimal control**
  - A popular framework: “predict-then-plan”

Motivation



- Single-agent **optimal control**
  - A popular framework: “predict-then-plan”
- Multi-agent **dynamic game**
  - tightly **coupled** plans
  - “**simultaneous** predict and plan”

Motivation



- Single-agent **optimal control**
  - A popular framework: “predict-then-plan”
- Multi-agent **dynamic game**
  - tightly **coupled** plans
  - “**simultaneous** predict and plan”

TU Delft

"Forward" Dynamic Games

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

other players' trajectories

$$\forall i \in [N] \begin{cases} \min\limits_{X^i, U^i} & J^i(X^i, U^i, \mathbf{X}^{\neg i}, \mathbf{U}^{\neg i}; \theta^i) \\ \text{s.t.} & h^i(X^i, U^i) = 0 \\ & {}^p g^i(X^i, U^i) \geq 0 \\ & {}^s g(X^i, U^i, \mathbf{X}^{\neg i}, \mathbf{U}^{\neg i}) \geq 0 \end{cases}$$

objective: e.g., drive fast

e.g., vehicle dynamics

e.g., max. speed

e.g., collision avoidance (shared)

Solution: **generalized Nash equilibrium (GNE)**

$$(\mathbf{X}^*, \mathbf{U}^*) := ((X^{1*}, U^{1*}), \ldots, (X^{N*}, U^{N*}))$$

TU Delft

"Forward" Dynamic Games

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

other players' trajectories

$$\forall i \in [N] \begin{cases} \min\limits_{X^i, U^i} & J^i(X^i, U^i, \mathbf{X}^{\neg i}, \mathbf{U}^{\neg i}; \theta^i) \\ \text{s.t.} & h^i(X^i, U^i) = 0 \\ & {}^p g^i(X^i, U^i) \geq 0 \\ & {}^s g(X^i, U^i, \mathbf{X}^{\neg i}, \mathbf{U}^{\neg i}) \geq 0 \end{cases}$$

objective: e.g., drive fast

e.g., vehicle dynamics

e.g., max. speed

e.g., collision avoidance (shared)

Solution: **generalized Nash equilibrium (GNE)**

Model-predictive game-play (MPGP): receding horizon

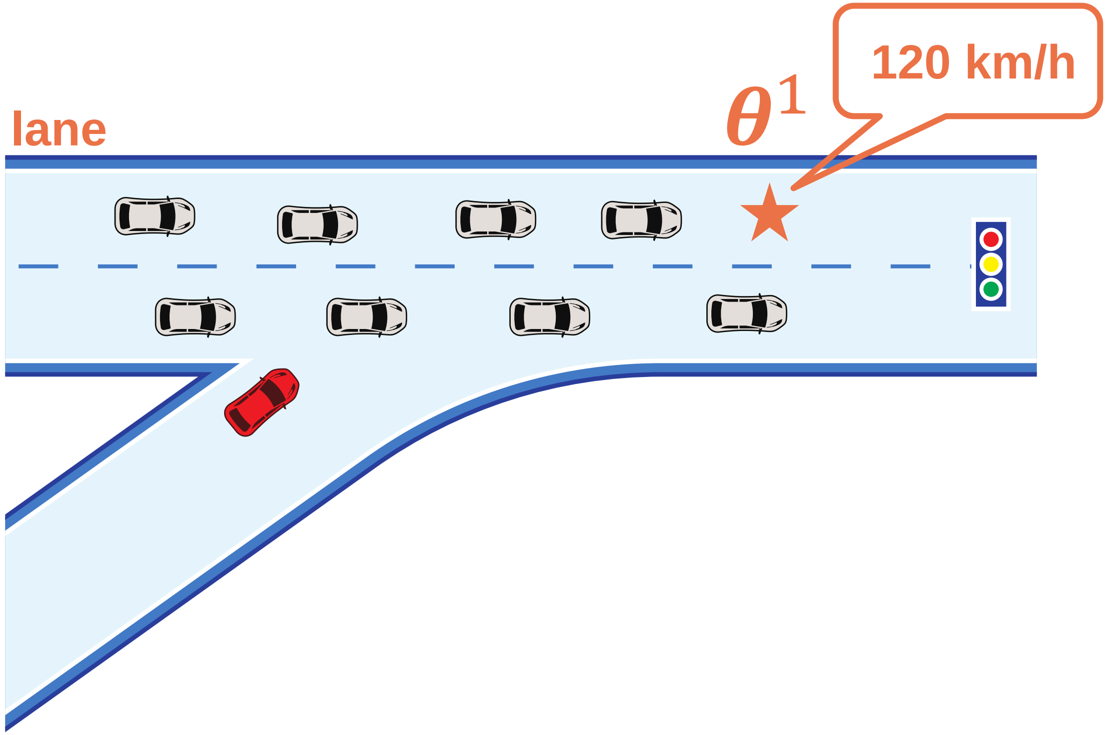
TU Delft

“Forward” Dynamic Games

$\boldsymbol{\theta}^{\neg 1}$? Unknown

e.g., desired speed, target lane

$$\forall i \in [N] \quad \begin{cases} \min_{X^i, U^i} & J^i(X^i, U^i, \mathbf{X}^{\neg i}, \mathbf{U}^{\neg i}; \theta^i) \\ \text{s.t.} & h^i(X^i, U^i) = 0 \\ & {}^p g^i(X^i, U^i) \geq 0 \\ & {}^s g(X^i, U^i, \mathbf{X}^{\neg i}, \mathbf{U}^{\neg i}) \geq 0 \end{cases}$$

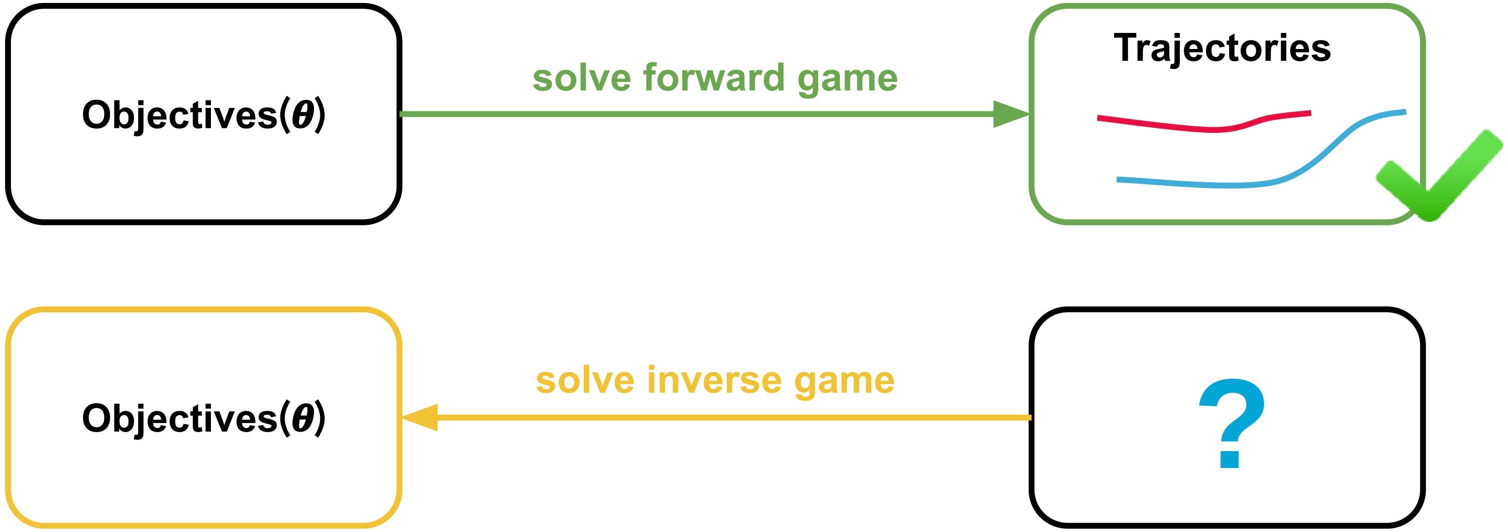


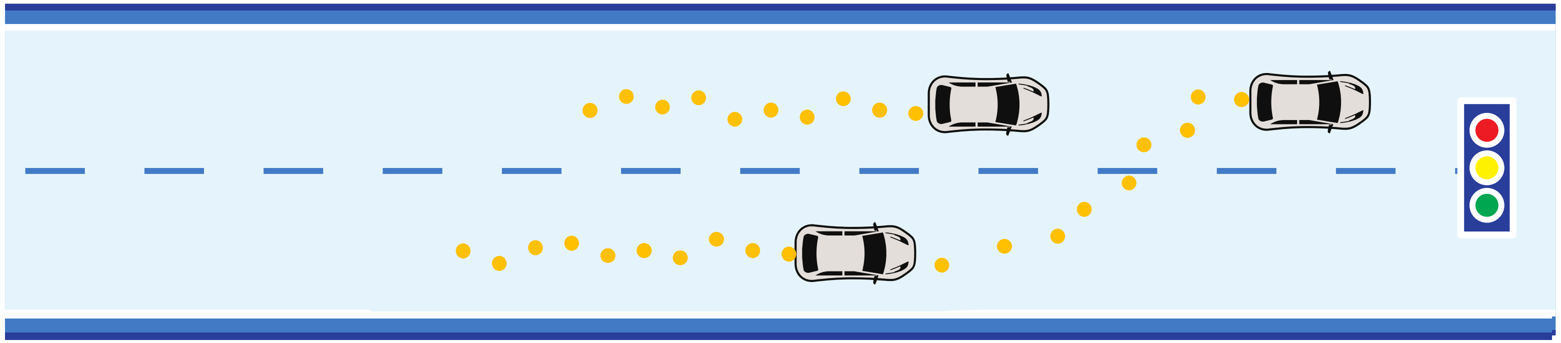
TU Delft

Model-predictive game-play (MPGP)
against opponents with **unknown** objectives ($\boldsymbol{\theta}$)

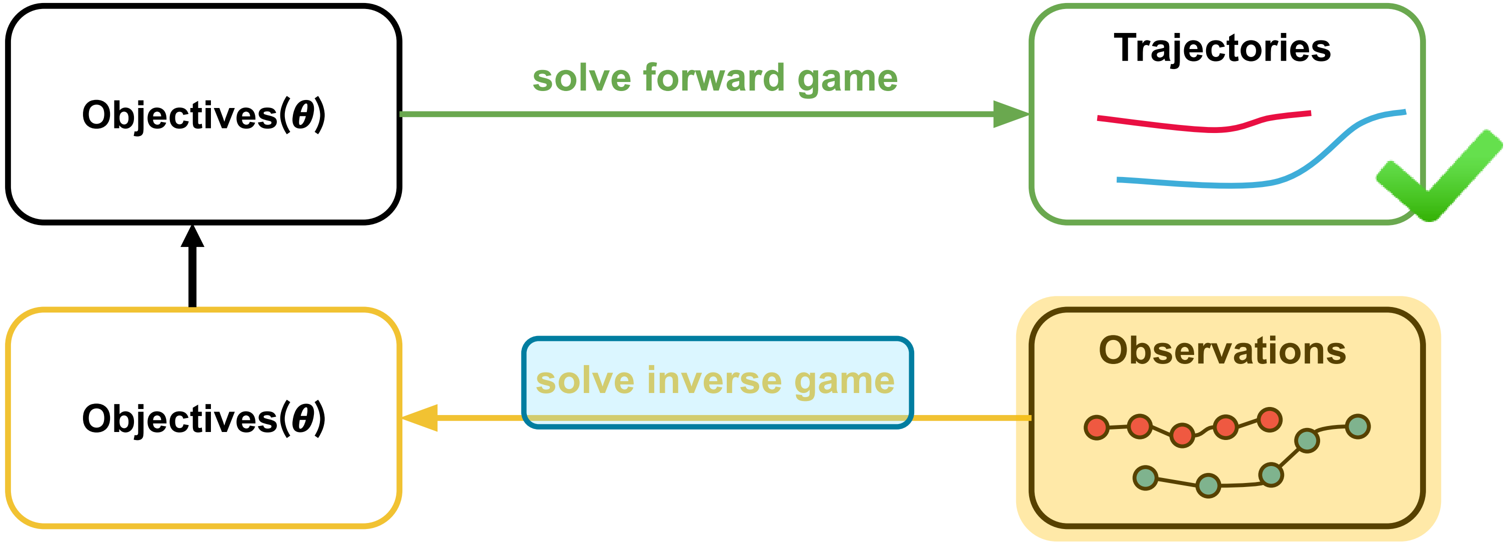
Forward and Inverse Games

Forward and Inverse Games

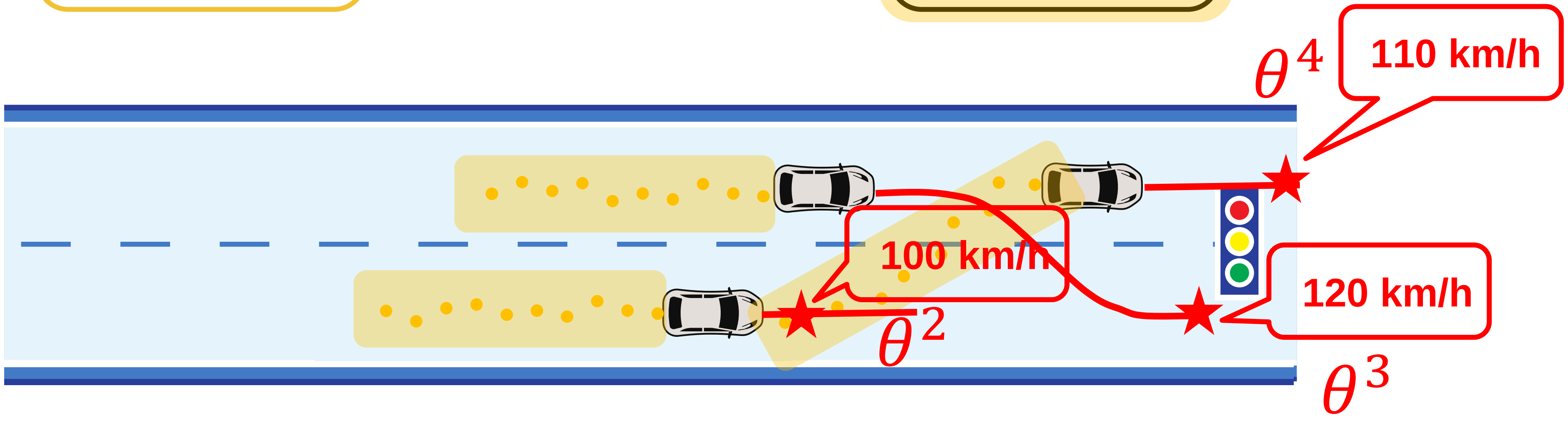




Forward and Inverse Games



Adaptive MPGP!



Inverse Games: Constrained Maximum Likelihood Estimation (MLE)

Observations Trajectory

$$\max_{\boldsymbol{\theta}, \mathbf{X}, \mathbf{U}} \quad p(\mathbf{Y} \mid \mathbf{X}, \mathbf{U})$$

e.g., opponents' desired speed, target lane

$$\text{s.t.} \quad (\mathbf{X}, \mathbf{U}) \text{ is a GNE of Game}(\boldsymbol{\theta})$$

Inverse Games: Constrained Maximum Likelihood Estimation (MLE)

$$\max_{\boldsymbol{\theta}, \mathbf{X}, \mathbf{U}} \quad p(\mathbf{Y} \mid \mathbf{X}, \mathbf{U})$$

$$\text{s.t.} \quad (\mathbf{X}, \mathbf{U}) \text{ is a GNE of Game}(\boldsymbol{\theta})$$

optimality (KKT) conditions of a forward game

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

$$h(\mathbf{X}, \mathbf{U}; \hat{\mathbf{x}}_1) = 0$$

$$0 \leq {}^s g(\mathbf{X}, \mathbf{U}) \perp {}^s\lambda \geq 0$$

Inverse Games: Constrained Maximum Likelihood Estimation (MLE)

$$\max_{\boldsymbol{\theta}, \mathbf{X}, \mathbf{U}} \quad p(\mathbf{Y} \mid \mathbf{X}, \mathbf{U})$$
$$\text{s.t.} \quad (\mathbf{X}, \mathbf{U}) \text{ is a GNE of Game}(\boldsymbol{\theta})$$

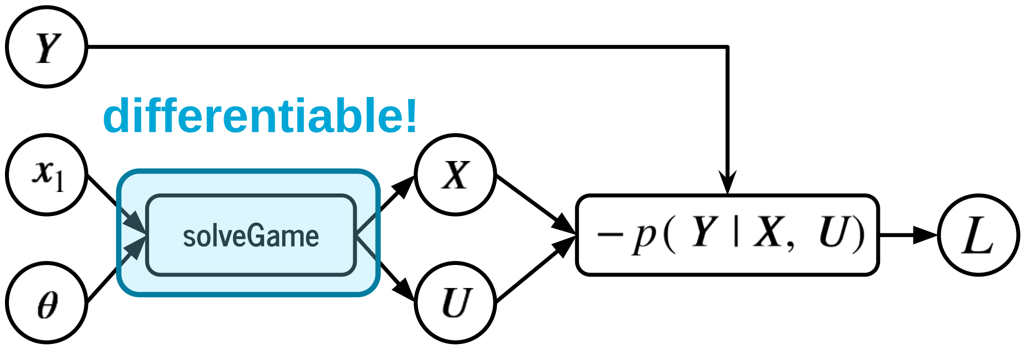
Challenge: how to efficiently encode the **equilibrium constraints**?

- Nonconvexity
- Complementarity conditions ✗constraint qualification
- Real-time computation

TU Delft

Approach: Differentiable Games

The Forward Computation Graph

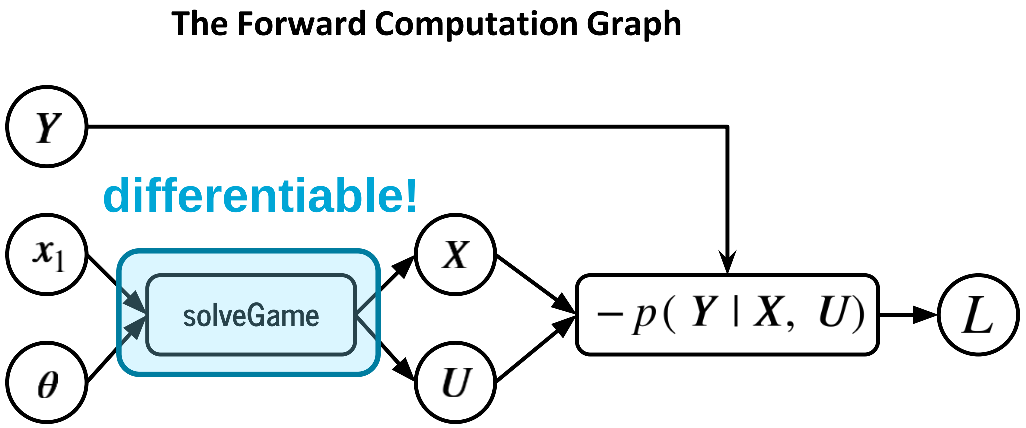


This entire computation graph can be made **differentiable**!

$$\max_{\boldsymbol{\theta}, \mathbf{X}, \mathbf{U}} \quad p(\mathbf{Y} \mid \mathbf{X}, \mathbf{U})$$

$$\text{s.t.} \quad (\mathbf{X}, \mathbf{U}) \text{ is a GNE of Game}(\boldsymbol{\theta})$$

Approach: Differentiable Games



This entire computation graph can be made **differentiable**!

⇒ We can update estimates of $\boldsymbol{\theta}$ via **gradient descent** on the loss function.

How?

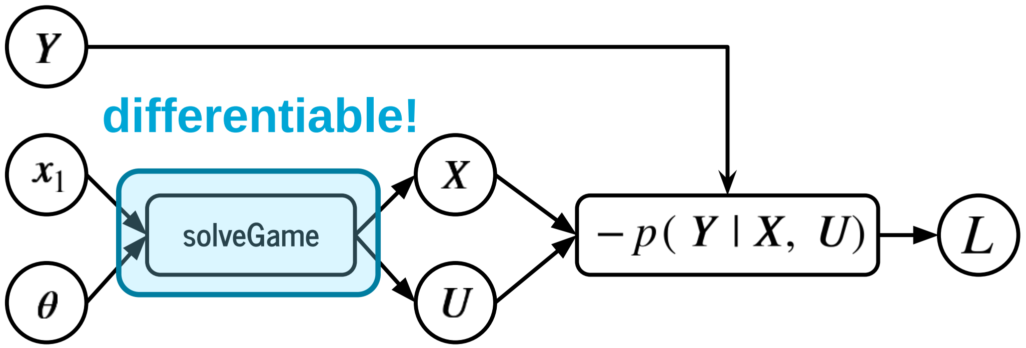
$$\max_{\boldsymbol{\theta}} \quad p(\mathbf{Y} \mid \mathbf{X}^*(\boldsymbol{\theta}), \mathbf{U}^*(\boldsymbol{\theta}))$$

✔ $\nabla_{\boldsymbol{\theta}} p!$

TU Delft

Approach: Differentiable Games

The Forward Computation Graph



This entire computation graph can be made **differentiable**!

⇒ We can update estimates of $\boldsymbol{\theta}$ via **gradient descent** on the loss function.

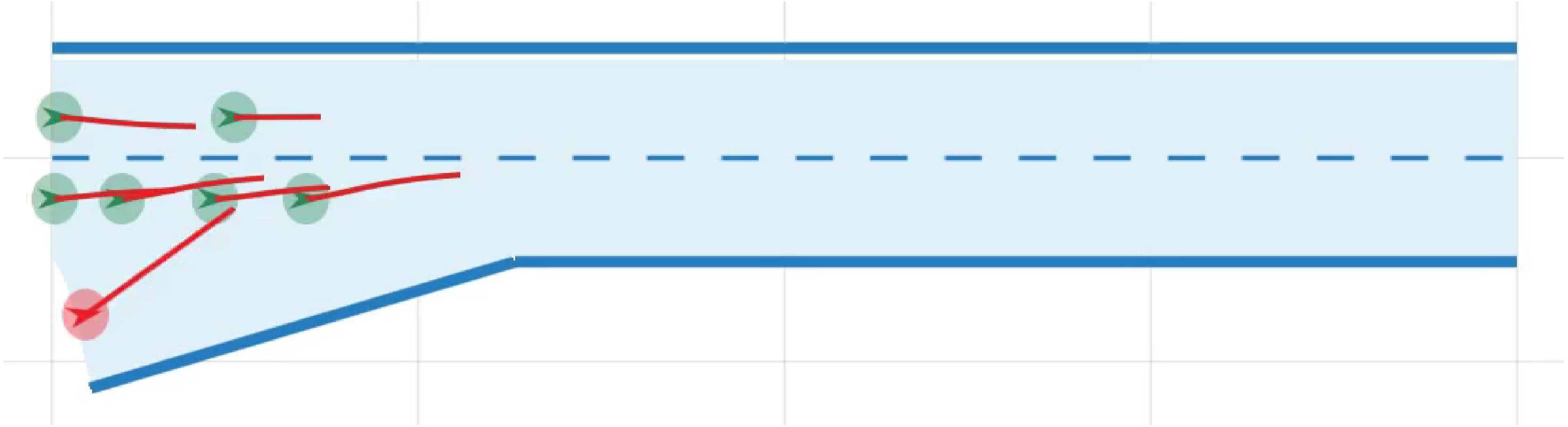
How?

Math! (Implicit function theorem)

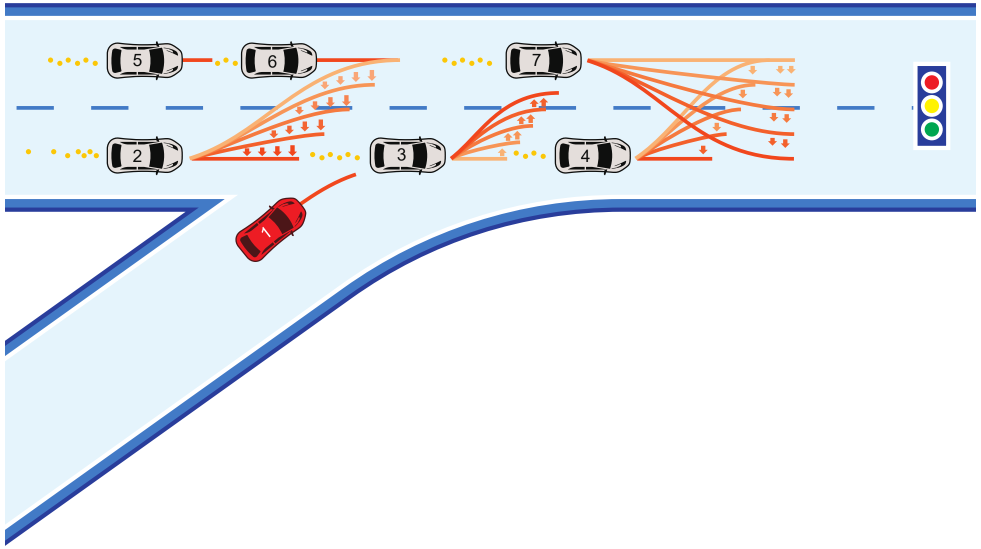
TU Delft

Results

Example: Ramp-Merging



Example: Ramp-Merging

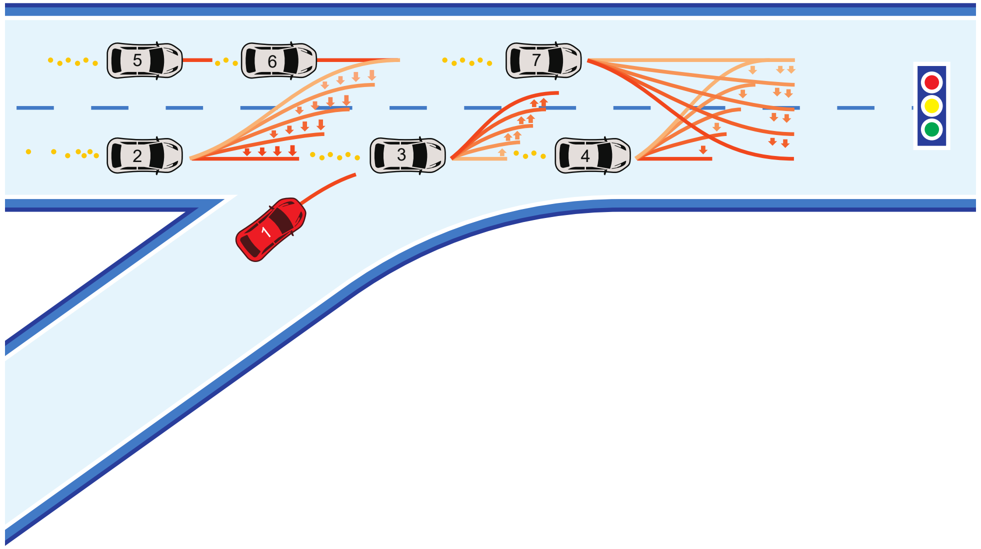


- Monte Carlo study: 3, 5, 7 players; 1200 trials total
- Unknown parameters: desired speed, target lane (12D for 6 opponents)
- Approaches

	Inverse game		Forward game
	Objective inference	Collis. avoid. inequalities	
Ours	✓	✓	✓

TU Delft

Example: Ramp-Merging



- Monte Carlo study: 3, 5, 7 players; 1200 trials total
- Unknown parameters: desired speed, target lane (12D for 6 opponents)
- Approaches

	Inverse game		Forward game
	Objective inference	Collis. avoid. inequalities	
Ours	✓	✓	✓
KKT-constrained [1]	✓	✗	✓
Heuristic	✗	✗	✓
MPC	✗	✗	✗

Collis. inequalities (inverse game)

Objective inference

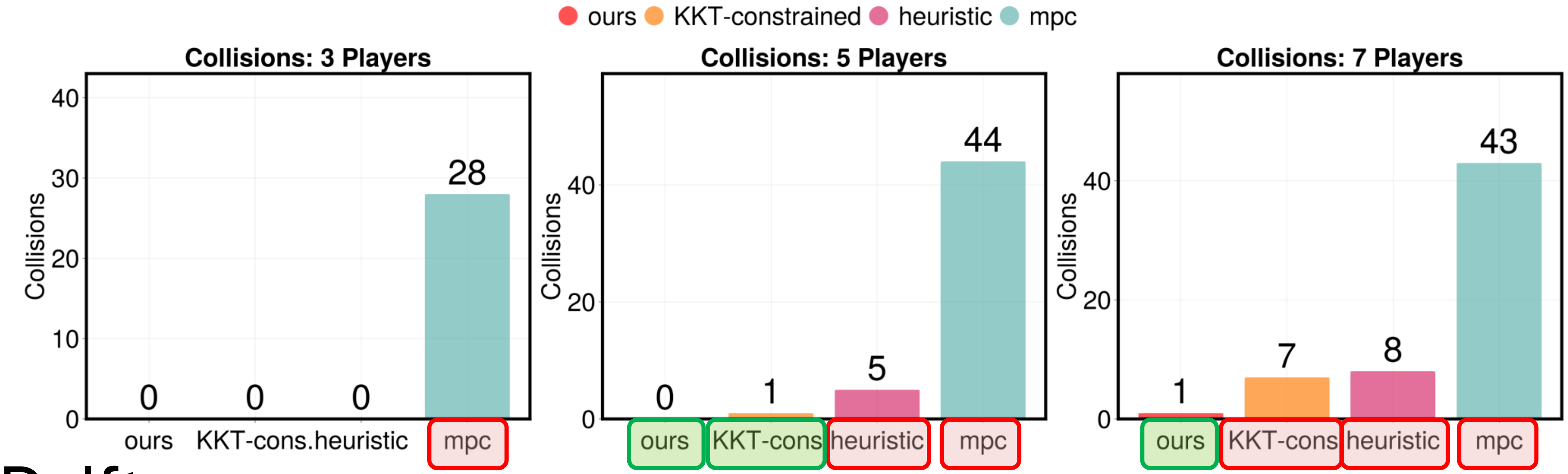
Interaction

TU Delft

[1] Lasse Peters et al. “Inferring Objectives in Continuous Dynamic Games from Noise-Corrupted Partial State Observations”. RSS. 2021.

Ramp-Merging: Quantitative Results

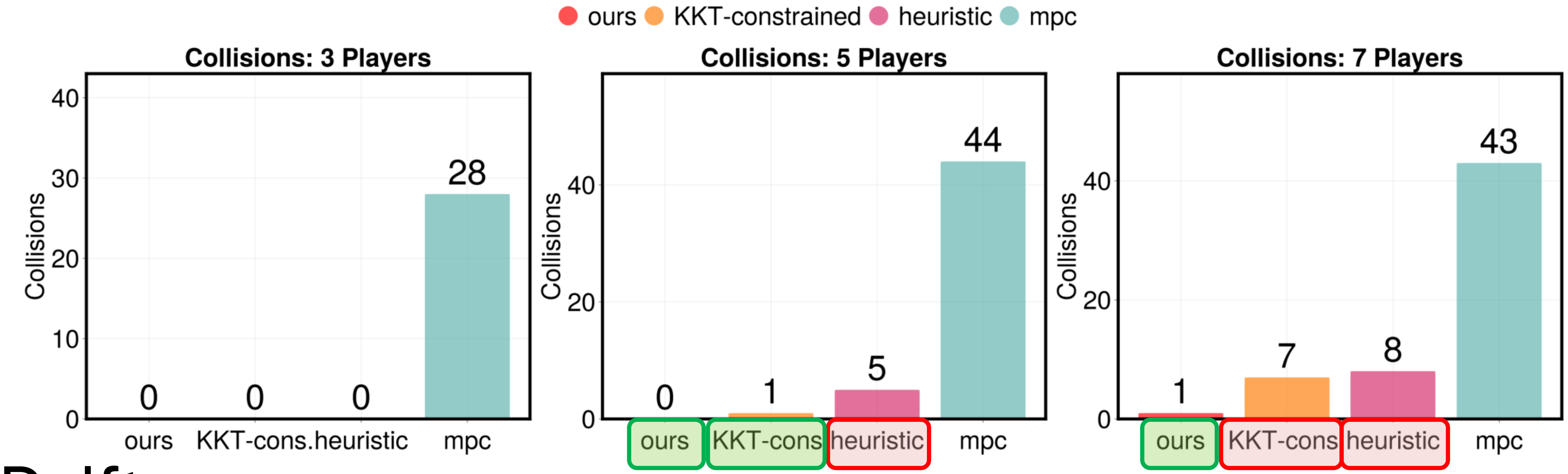
- Safety (fewer collisions)
  - **Interaction** reasoning is essential



TUDelft

Ramp-Merging: Quantitative Results

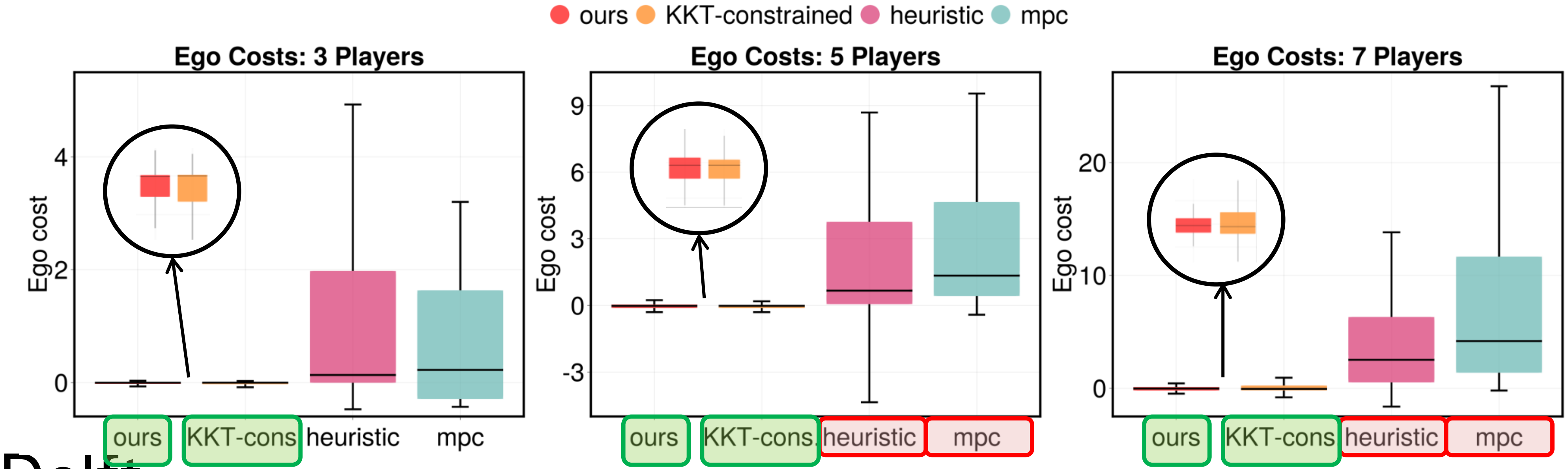
- Safety (fewer collisions)
  - Interaction reasoning is essential
  - Care about objective inference and inverse game inequalities in dense scenarios



Ramp-Merging: Quantitative Results

- Safety (fewer collisions)
  - Interaction reasoning is essential
  - Care about objective inference and inverse game inequalities in dense scenarios

- Efficiency (lower ego costs)
  - Interaction reasoning & objective inference are important in denser settings
  - Collision avoidance inequalities in inverse games does not matter



TU Delft

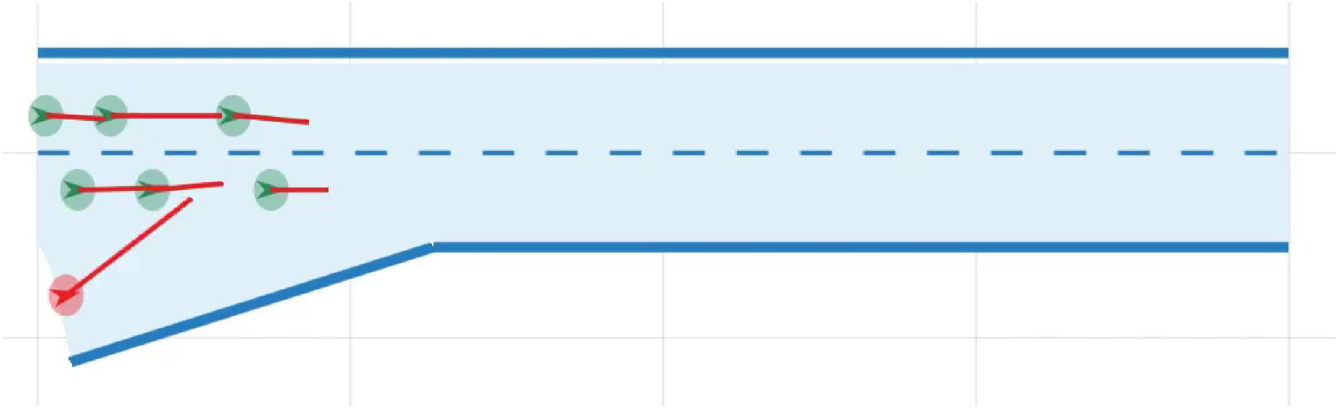
Ramp-Merging: Conclusions

- Safety (fewer collisions)
  - Interaction reasoning is essential
  - Care about objective inference and inverse game inequalities in dense scenarios

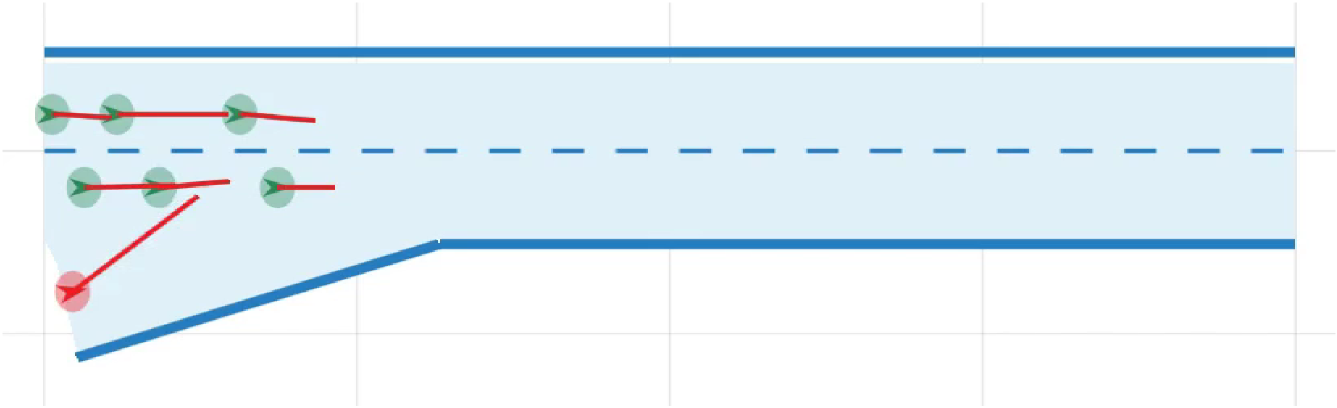
- Efficiency (lower ego costs)
  - Interaction reasoning & objective inference are important in denser settings
  - Collision avoidance inequalities in inverse games does not matter

Ramp-Merging: Qualitative Results

Ours

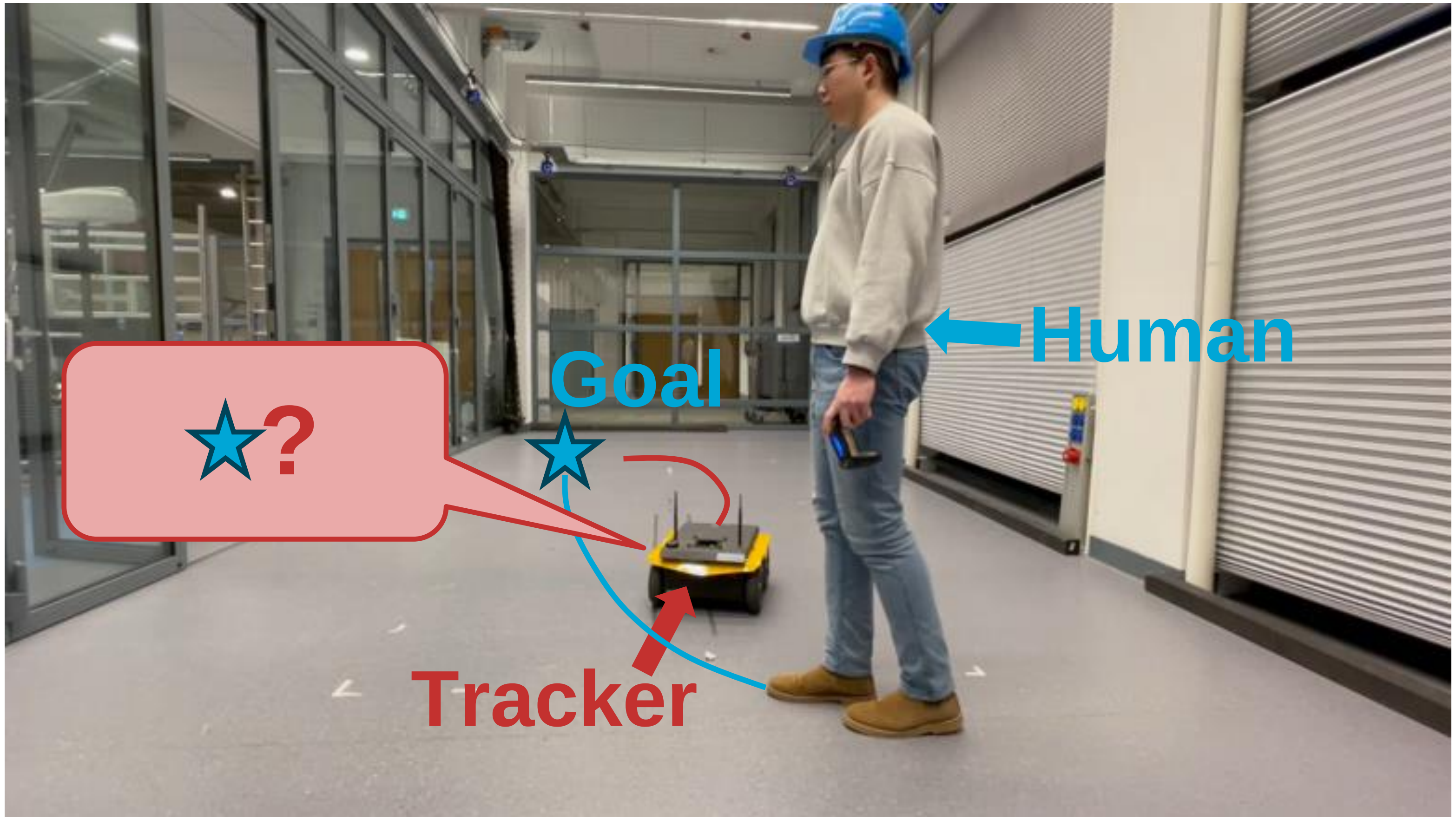


MPC



TUDelft

Example: 2-Player Tracking Game



TU Delft

More videos: https://www.youtube.com/watch?v=f0KJuCC1Xyo

Example: 2-Player Tracking Game

More videos: https://www.youtube.com/watch?v=f0KJuCC1Xyo

Summary

- An **adaptive model-predictive game-play (MPGP) framework** enabled by differentiating through a game solver
  - handling **inequalities** in inverse games
  - **differentiability**

Future Work

- **Planning algorithm** utilizing the beliefs (stochastic games)
- End-to-end planning pipeline with **perception module**

TU Delft

#betterTUgether

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 an open-loop game can be cast as the problem data of an equivalent MCP!

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

$\longrightarrow$

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

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

$\boldsymbol{\theta}$ → ***solveMCP*** → $\boldsymbol{z}^*$

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:

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

TU Delft

Approach: Differentiation Through Mixed Complementarity Problems

$\boldsymbol{\theta} \rightarrow$ ***solveMCP*** $\rightarrow \boldsymbol{z}^*$

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:

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

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

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

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

$$\tilde{\mathcal{I}} := \{k \in [n] \mid z_k^* = \ell_k \vee z_k^* = u_k\}$$

$$\tilde{z}^* := [\boldsymbol{z}^*]_{\tilde{\mathcal{I}}}$$

TU Delft

Approach: Differentiation Through Mixed Complementarity Problems

$\boldsymbol{\theta} \rightarrow$ ***solveMCP*** $\rightarrow \boldsymbol{z}^*$

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:

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

$$\ell_j < z_j^* < u_j, F_j(\boldsymbol{z}^*; \boldsymbol{\theta}) = 0 \qquad \nabla_{\boldsymbol{\theta}} \bar{z}^* = -\left(\nabla_{\bar{z}^*} \bar{F}\right)^{-1} \left(\nabla_{\boldsymbol{\theta}} \bar{F}\right)$$

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

$$\bar{\mathcal{I}} := \{k \in [n] \mid F_k(\boldsymbol{z}^*; \boldsymbol{\theta}) = 0, \ell_k < z_k^* < u_k\}, \quad \bar{z}^* := [\boldsymbol{z}^*]_{\bar{\mathcal{I}}}, \quad \bar{F}(\boldsymbol{z}^*, \boldsymbol{\theta}) := [F(\boldsymbol{z}^*; \boldsymbol{\theta})]_{\bar{\mathcal{I}}}$$

Implicit function theorem (IFT):

if invertible

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

TU Delft

Approach: Differentiation Through Mixed Complementarity Problems

$\boldsymbol{\theta} \rightarrow$ **solveMCP** $\rightarrow \boldsymbol{z}^*$

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:

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

$$\ell_j < z_j^* < u_j, F_j(\boldsymbol{z}^*; \boldsymbol{\theta}) = 0 \qquad \nabla_{\boldsymbol{\theta}} \bar{z}^* = -\left(\nabla_{\bar{z}^*} \bar{F}\right)^{-1} \left(\nabla_{\boldsymbol{\theta}} \bar{F}\right)$$

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

$$\bar{\mathcal{I}} := \{k \in [n] \mid F_k(\boldsymbol{z}^*; \boldsymbol{\theta}) = 0, \ell_k < z_k^* < u_k\}, \quad \bar{z}^* := [\boldsymbol{z}^*]_{\bar{\mathcal{I}}}, \quad \bar{F}(\boldsymbol{z}^*, \boldsymbol{\theta}) := [F(\boldsymbol{z}^*; \boldsymbol{\theta})]_{\bar{\mathcal{I}}}$$

Implicit function theorem (IFT):

if invertible

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