Imitative Planning

for Autonomous Vehicles

Nicholas Rhinehart
Carnegie Mellon University

Rowan McAllister
UC Berkeley

Kris Kitani
Carnegie Mellon University

Sergey Levine
UC Berkeley

CARLA

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

Without learning ?

Behavior cloning ?

Model-based RL ?

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

(trains offline)

Without learning ?

Behavior cloning ?

Model-based RL ?

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

(trains offline)

Without learning ? ✔

Behavior cloning ? ?

Model-based RL ? ?

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

(trains offline) (no rewards required)

Without learning ? ✔

Behavior cloning ? ?

Model-based RL ? ?

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

(trains offline) (no rewards required)

Without learning ?	✓	✗
Behavior cloning ?	?	✓
Model-based RL ?	?	✗

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

(trains offline) (no rewards required) (plans)

Without learning ?	✓	✗
Behavior cloning ?	?	✓
Model-based RL ?	?	✗

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

	(trains offline)	(no rewards required)	(plans)
Without learning ?	✓	✗	✓
Behavior cloning ?	?	✓	✗
Model-based RL ?	?	✗	✓

Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

	(trains offline)	(no rewards required)	(plans)	(dyn. model)
Without learning ?	✓	✗	✓	
Behavior cloning ?	?	✓	✗	
Model-based RL ?	?	✗	✓	

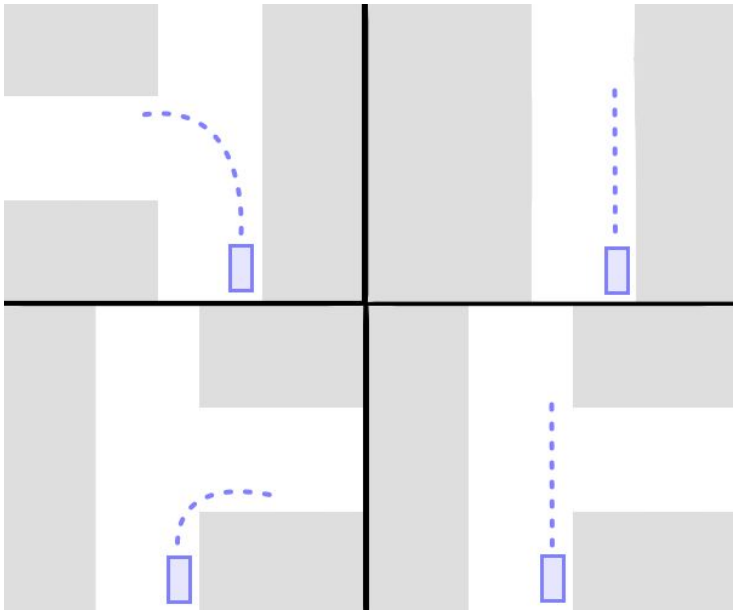
Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

	(trains offline)	(no rewards required)	(plans)	(dyn. model)
Without learning ?	✓	✗	✓	?
Behavior cloning ?	?	✓	✗	✗
Model-based RL ?	?	✗	✓	✓

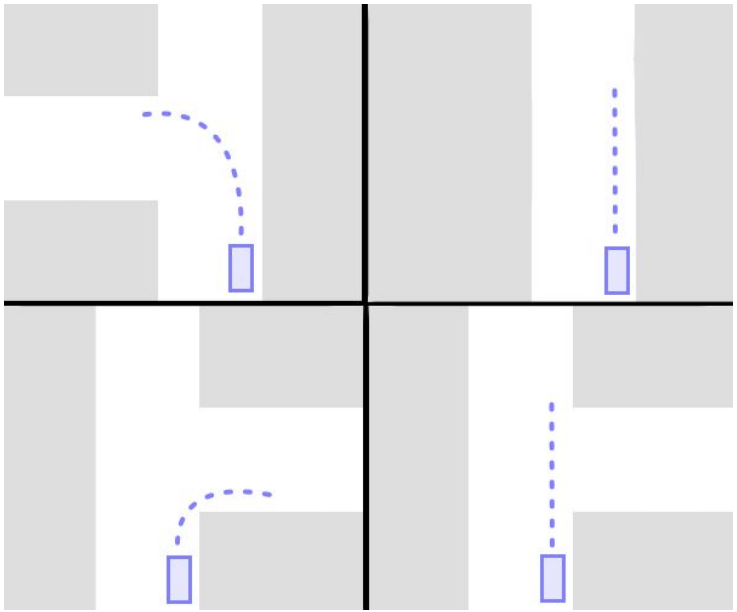
Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

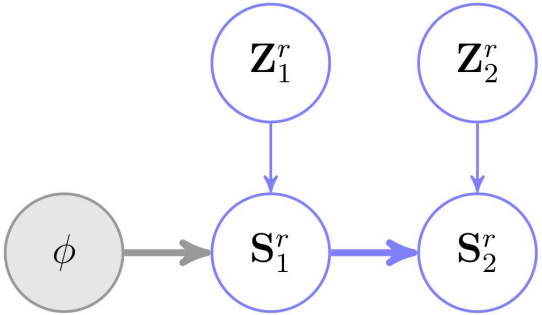
	(trains offline)	(no rewards required)	(plans)	(dyn. model)
Without learning ?	✓	✗	✓	?
Behavior cloning ?	?	✓	✗	✗
Model-based RL ?	?	✗	✓	✓
Imitative Models (ours)	✓	✓	✓	✓

Modelling Expert Drivers

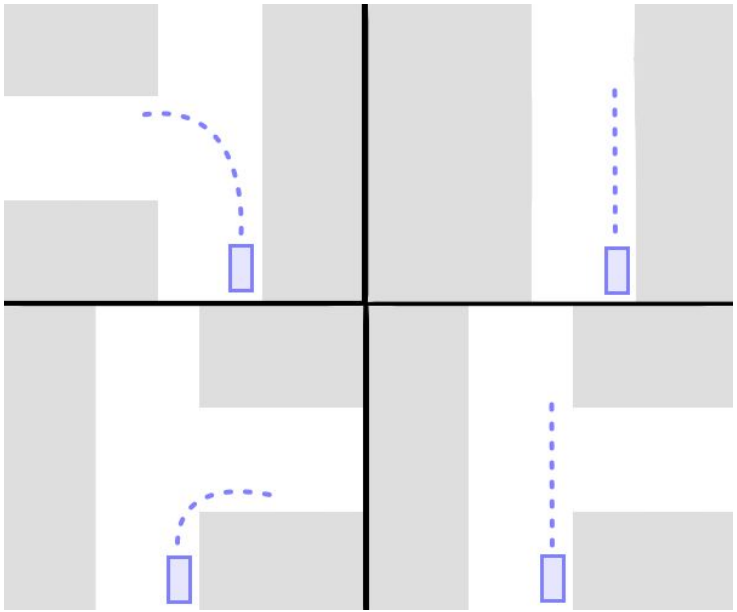


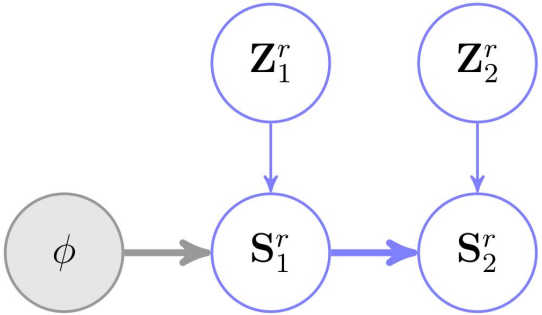
Modelling Expert Drivers





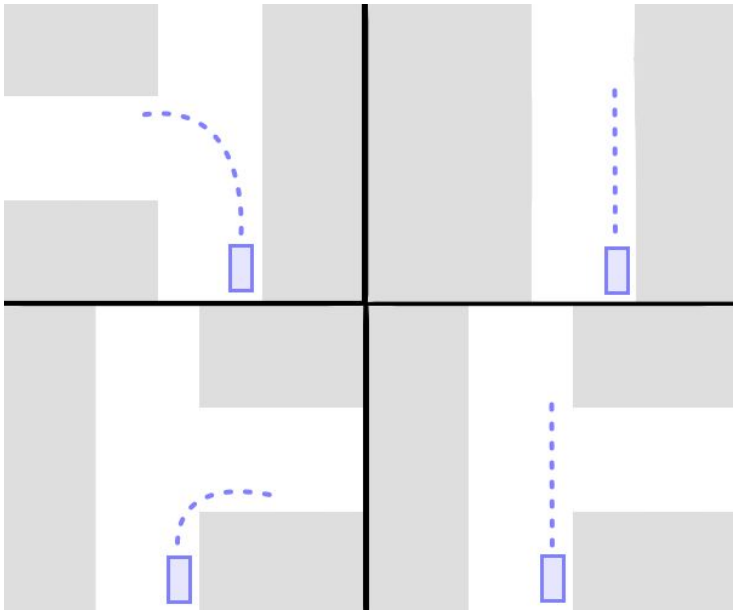
Modelling Expert Drivers

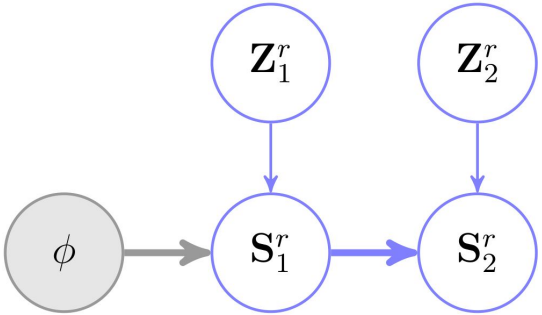




$$\textbf{states} = f\,(\textbf{latents}\,;\textbf{context})$$

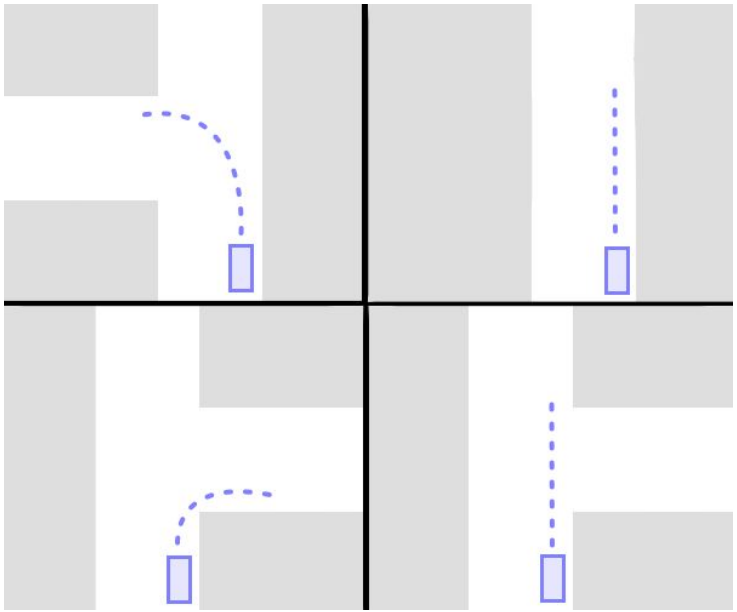
Modelling Expert Drivers

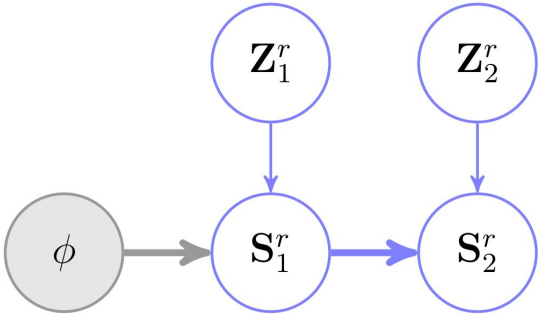




$$\text{states} = f(\text{latents}; \text{context})$$
$$\text{latents} = f^{-1}(\text{states}; \text{context})$$

Modelling Expert Drivers

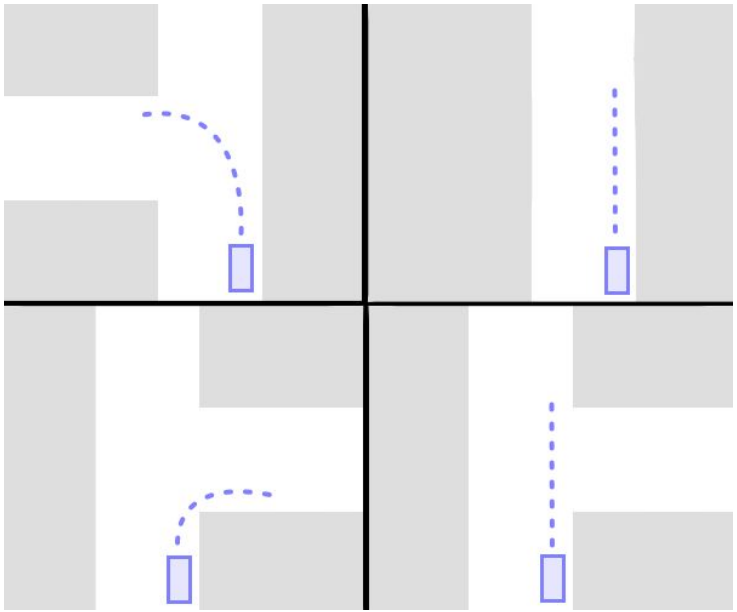


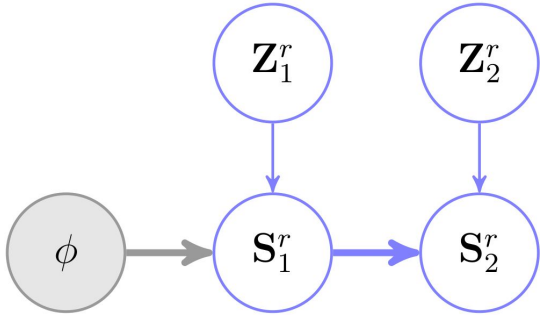


$$\text{states} = f(\text{latents}; \text{context})$$
$$\text{latents} = f^{-1}(\text{states}; \text{context})$$

$$q(\text{states} \mid \text{context}) = \frac{\mathcal{N}(\text{latents}; 0, I)}{\left|\det \frac{\partial f}{\partial \text{latents}}\right|}$$

Modelling Expert Drivers





$$\text{states} = f(\text{latents}; \text{context})$$
$$\text{latents} = f^{-1}(\text{states}; \text{context})$$

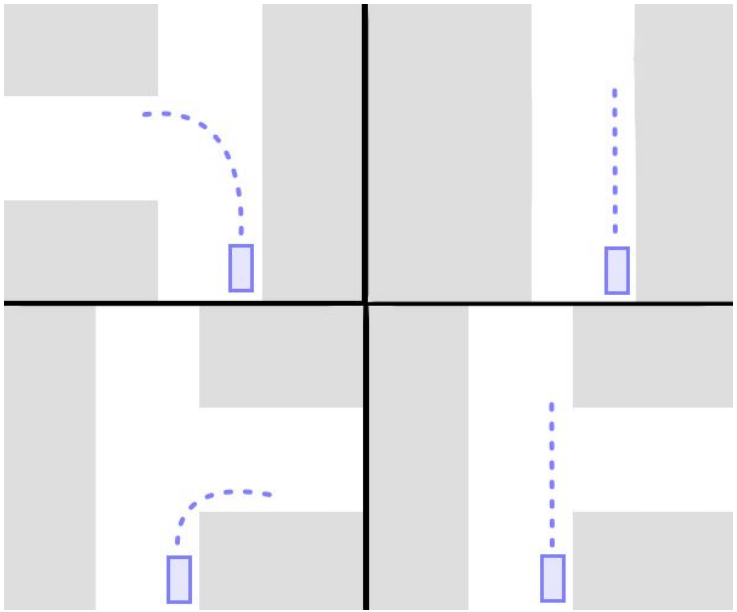
$$q(\text{states} \mid \text{context}) = \frac{\mathcal{N}(\text{latents}; 0, I)}{\left|\det \frac{\partial f}{\partial \text{latents}}\right|}$$

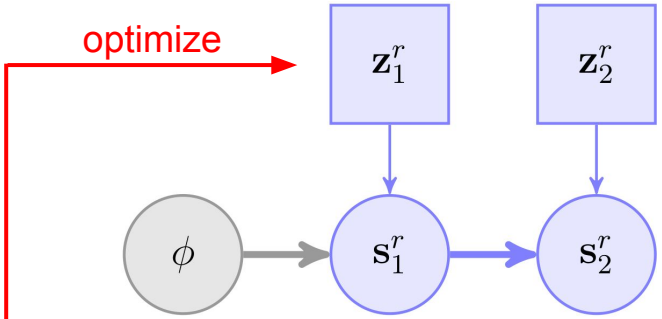
Can **plan** in this distribution!

new idea

Rhinehart, McAllister, Levine **http://imitate.ml**

Modelling Expert Drivers





$$\text{states} = f(\text{latents}; \text{context})$$
$$\text{latents} = f^{-1}(\text{states}; \text{context})$$

$$q(\text{states} \mid \text{context}) = \frac{\mathcal{N}(\text{latents}; 0, I)}{\left|\det \frac{\partial f}{\partial \text{latents}}\right|}$$

Can **plan** in this distribution!

new idea

Rhinehart, McAllister, Levine **http://imitate.ml**

Planning “Expert-like” Motions to our own Goals

$$\textbf{planned path} = \arg\max_{\textbf{states}} \log p(\textbf{states} \mid \textbf{goal}, \textbf{context})$$

Planning “Expert-like” Motions to our own Goals

$$\textbf{planned path} = \arg\max_{\textbf{states}} \log p(\textbf{states} \mid \textbf{goal}, \textbf{context})$$

$$= \arg\max_{\textbf{states}} \log q(\textbf{states} \mid \textbf{context}) + \log p(\textbf{goal} \mid \textbf{states}, \textbf{context}) - \log p(\textbf{goal} \mid \textbf{context})$$

Planning “Expert-like” Motions to our own Goals

$$
\begin{aligned}
\textbf{planned path} &= \arg\max_{\textbf{states}} \log p(\textbf{states} \mid \textbf{goal}, \textbf{context}) \\
&= \arg\max_{\textbf{states}} \log q(\textbf{states} \mid \textbf{context}) + \log p(\textbf{goal} \mid \textbf{states}, \textbf{context}) - \log p(\textbf{goal} \mid \textbf{context}) \\
&= \arg\max_{\textbf{states}} \log \underbrace{q(\textbf{states} \mid \textbf{context})}_{\text{imitation prior}} + \log \underbrace{p(\textbf{goal} \mid \textbf{states}, \textbf{context})}_{\text{goal likelihood}}
\end{aligned}
$$

Planning “Expert-like” Motions to our own Goals

$$
\begin{aligned}
\textbf{planned path} &= \arg\max_{\textbf{states}} \log p(\textbf{states} \mid \textbf{goal}, \textbf{context}) \\
&= \arg\max_{\textbf{states}} \log q(\textbf{states} \mid \textbf{context}) + \log p(\textbf{goal} \mid \textbf{states}, \textbf{context}) - \log p(\textbf{goal} \mid \textbf{context}) \\
&= \arg\max_{\textbf{states}} \log \underbrace{q(\textbf{states} \mid \textbf{context})}_{\text{imitation prior}} + \log \underbrace{p(\textbf{goal} \mid \textbf{states}, \textbf{context})}_{\text{goal likelihood}}
\end{aligned}
$$

Drive safely...

Rhinehart, McAllister, Levine **http://imitate.ml**

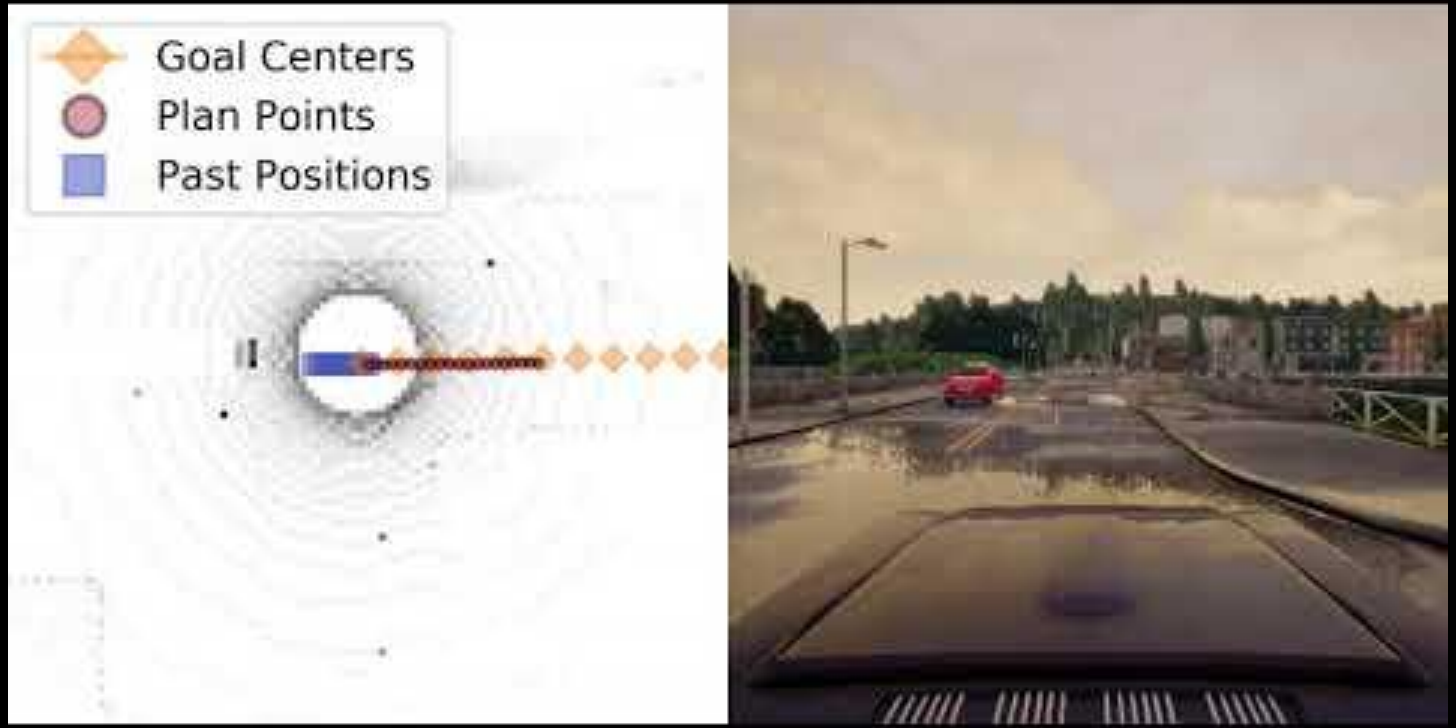
Planning “Expert-like” Motions to our own Goals

$$
\begin{aligned}
\textbf{planned path} &= \arg\max_{\textbf{states}} \log p(\textbf{states} \mid \textbf{goal}, \textbf{context}) \\
&= \arg\max_{\textbf{states}} \log q(\textbf{states} \mid \textbf{context}) + \log p(\textbf{goal} \mid \textbf{states}, \textbf{context}) - \log p(\textbf{goal} \mid \textbf{context}) \\
&= \arg\max_{\textbf{states}} \log \underbrace{q(\textbf{states} \mid \textbf{context})}_{\text{imitation prior}} + \log \underbrace{p(\textbf{goal} \mid \textbf{states}, \textbf{context})}_{\text{goal likelihood}}
\end{aligned}
$$

Drive safely...

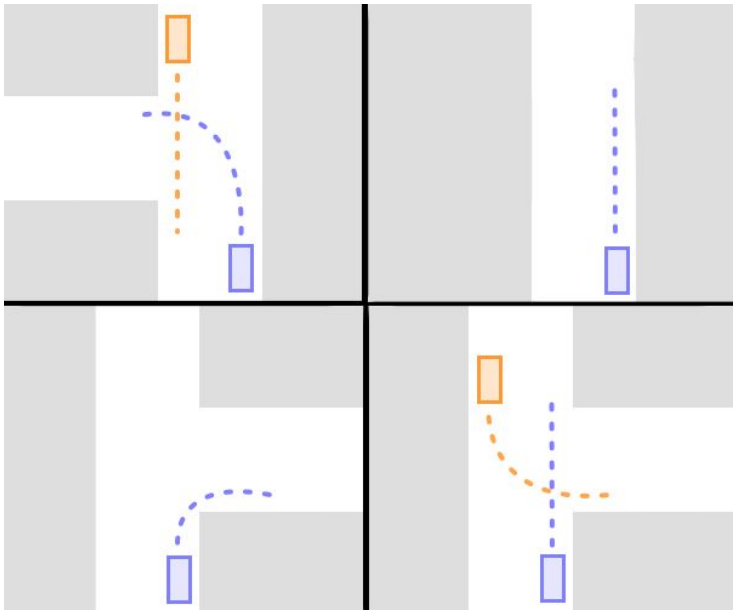
...and get to our destination!

Rhinehart, McAllister, Levine **http://imitate.ml**

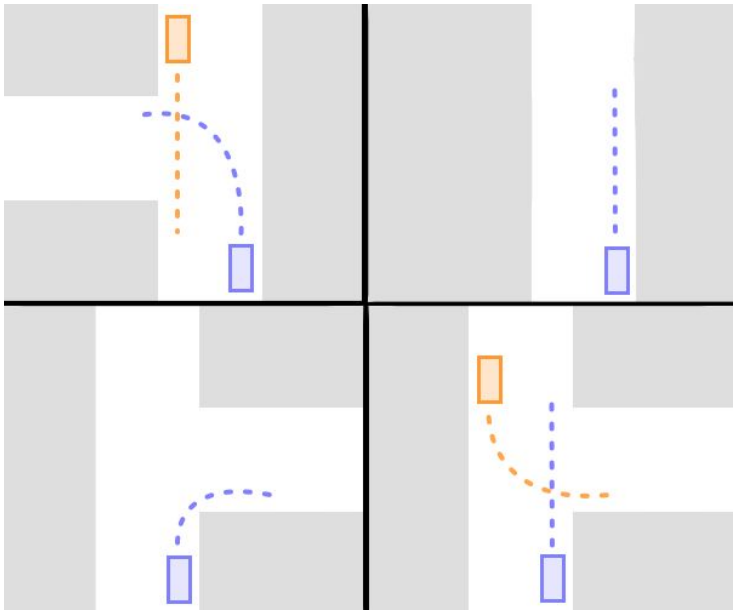

Goal Centers
Plan Points
Past Positions

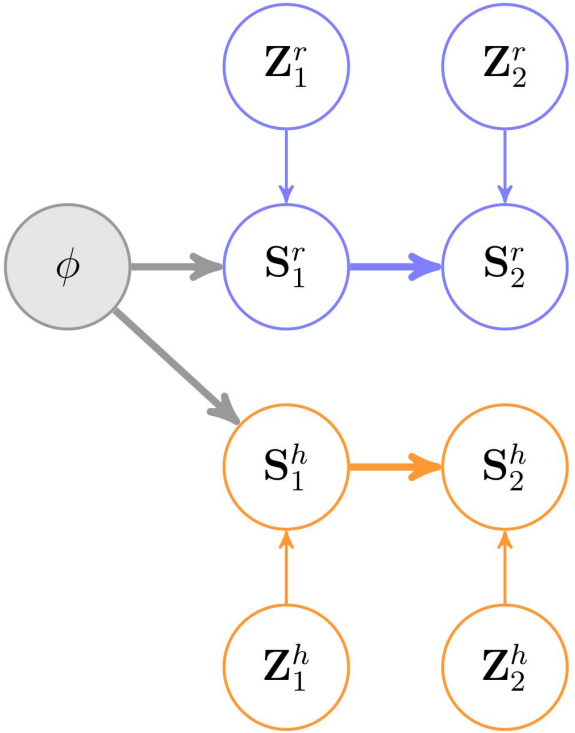
Goal Centers
Plan Points
Past Positions

Modelling Multiple Expert Drivers

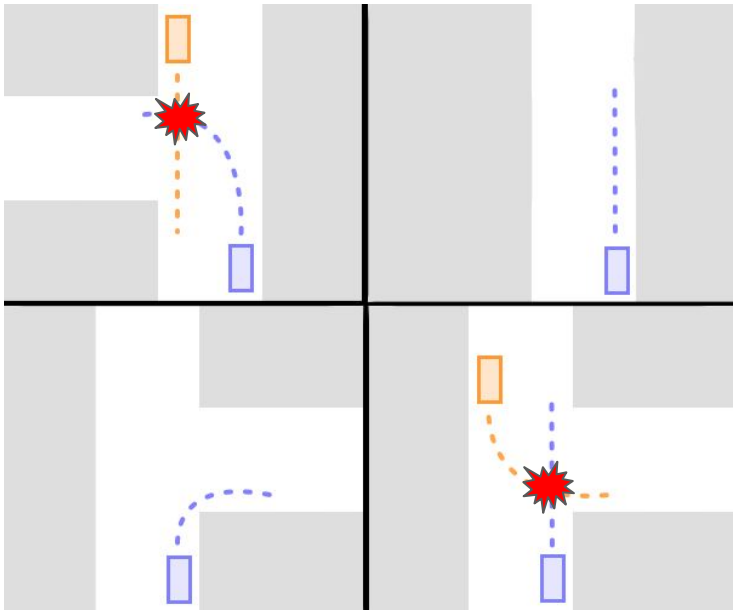


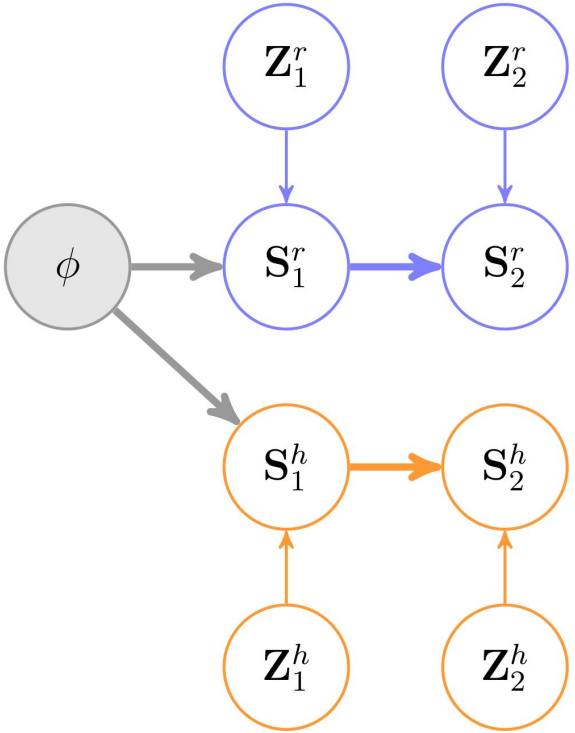
Modelling Multiple Expert Drivers



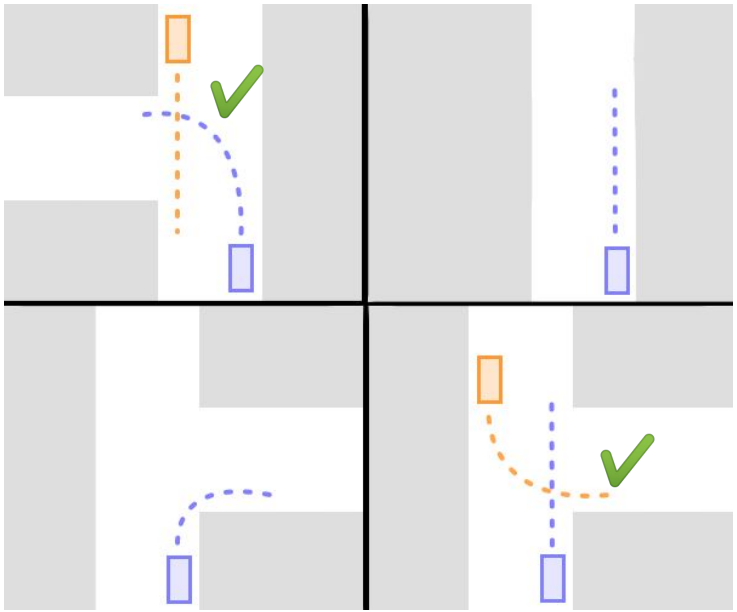


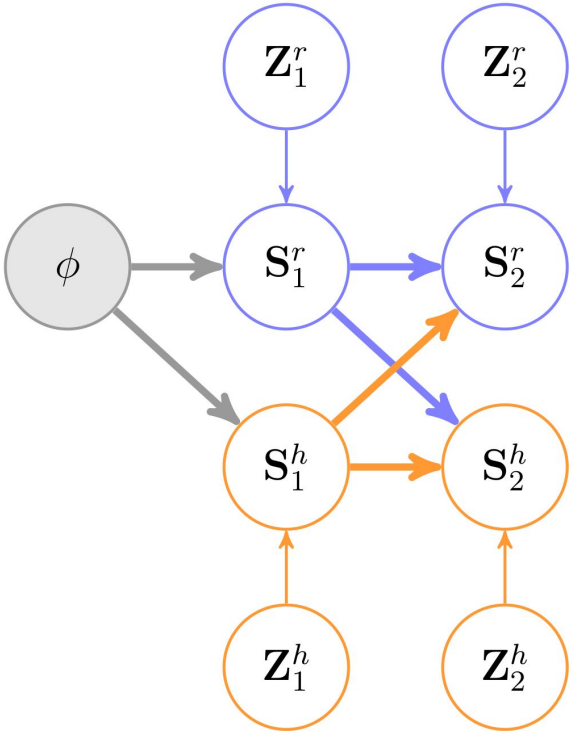
Modelling Multiple Expert Drivers



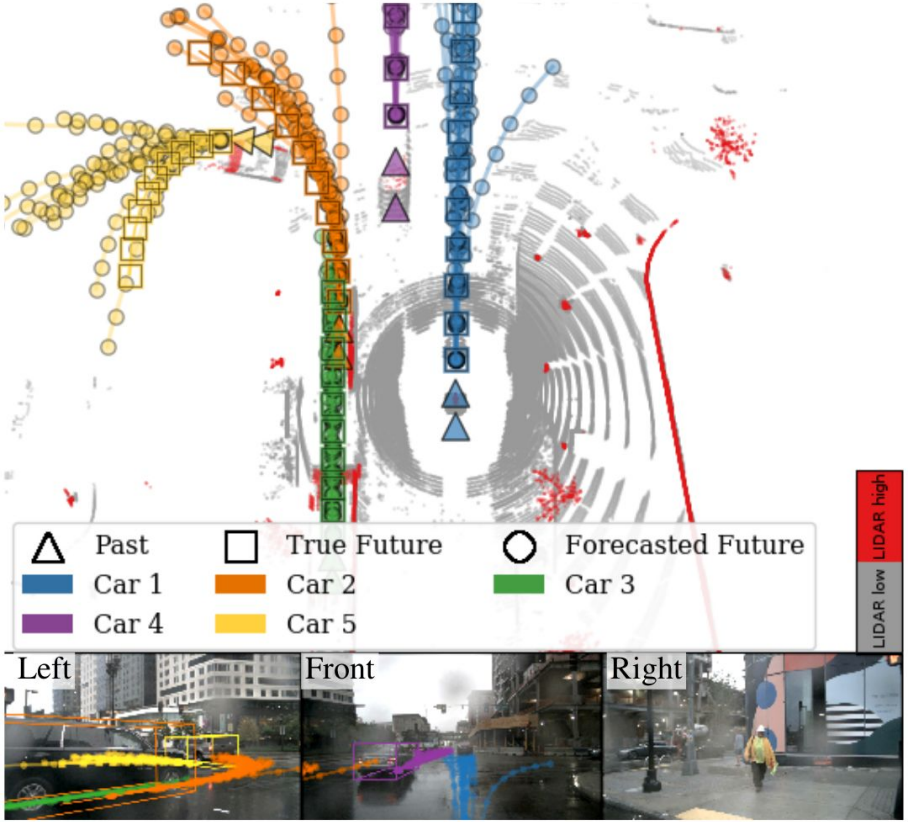


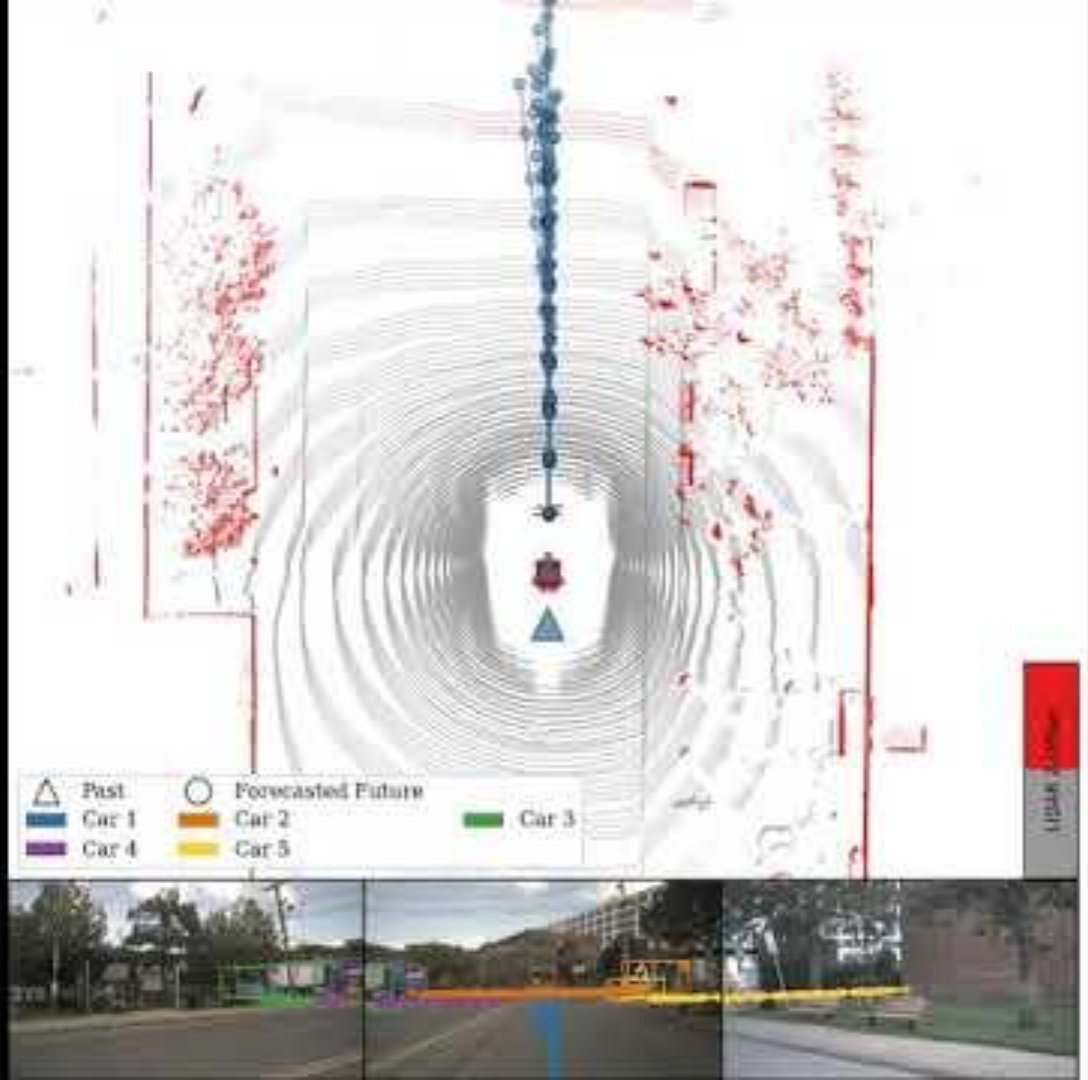
Modelling Multiple Expert Drivers





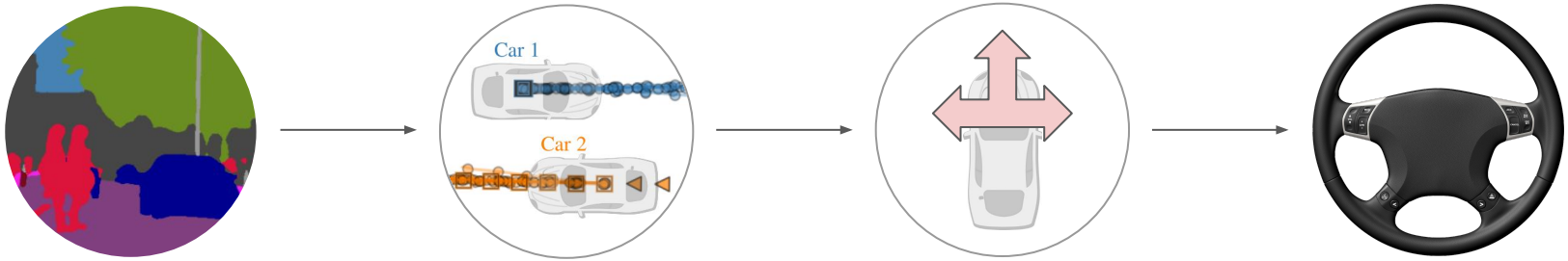
Forecasting with nuScenes data (Singapore + Boston)




Past
Forecasted Future
Car 1
Car 2
Car 3
Car 4
Car 5

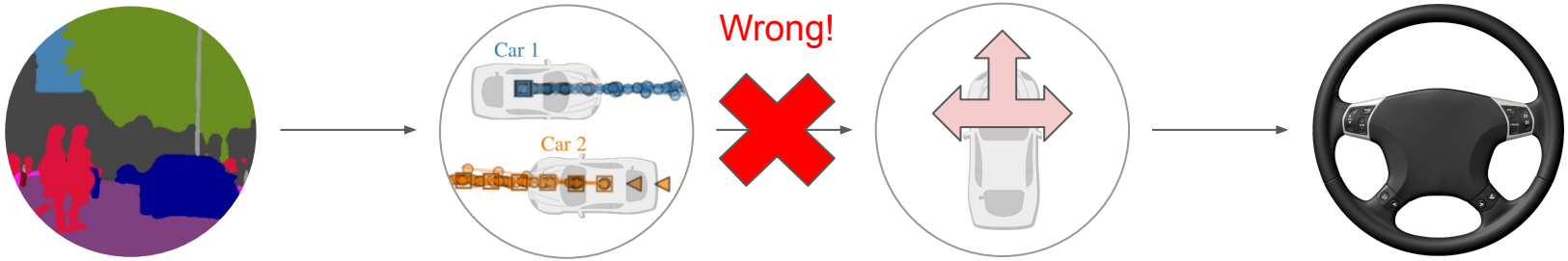
Multi-Agent Planning

Q: but how should ***autonomous vehicles*** predict other agents...**online**?



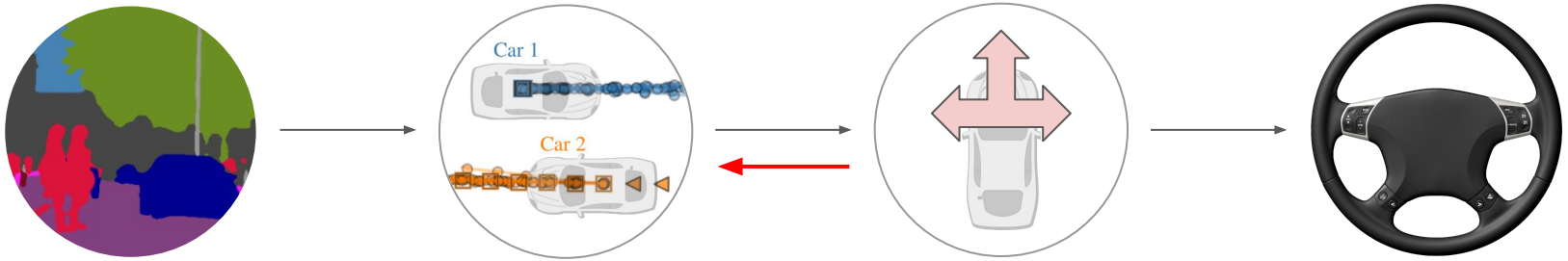
Multi-Agent Planning

Q: but how should ***autonomous vehicles*** predict other agents...**online**?

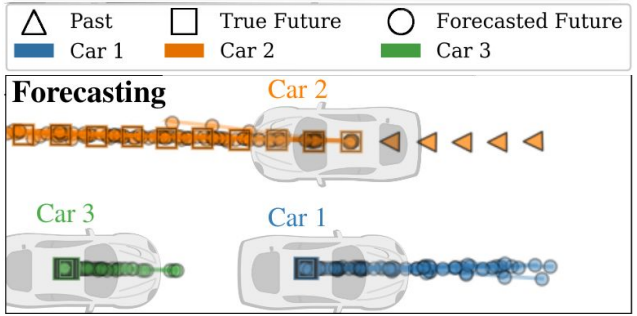


Multi-Agent Planning

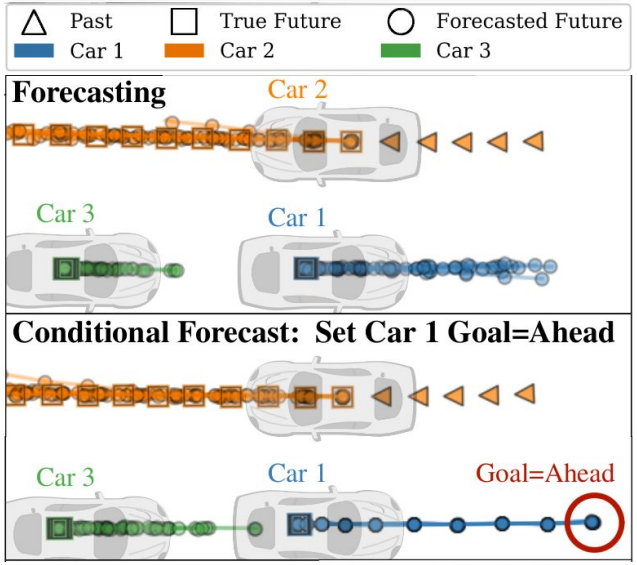
Q: but how should ***autonomous vehicles*** predict other agents...**online**?



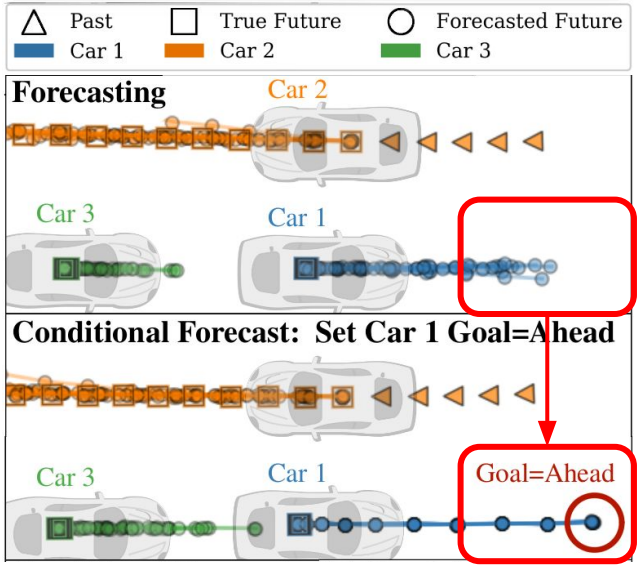
Goal-Conditioned Multi-Agent Forecasting



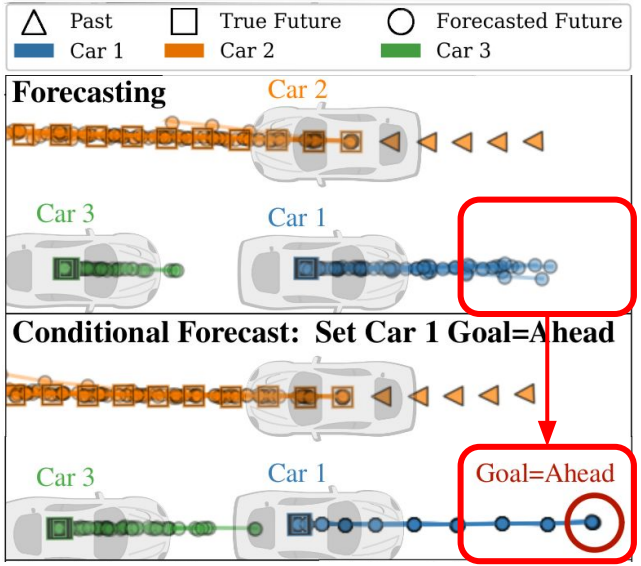
Goal-Conditioned Multi-Agent Forecasting



Goal-Conditioned Multi-Agent Forecasting

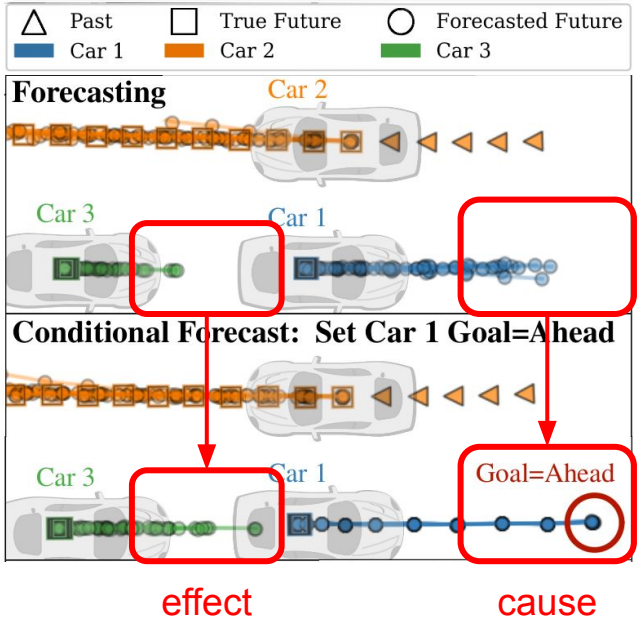


Goal-Conditioned Multi-Agent Forecasting

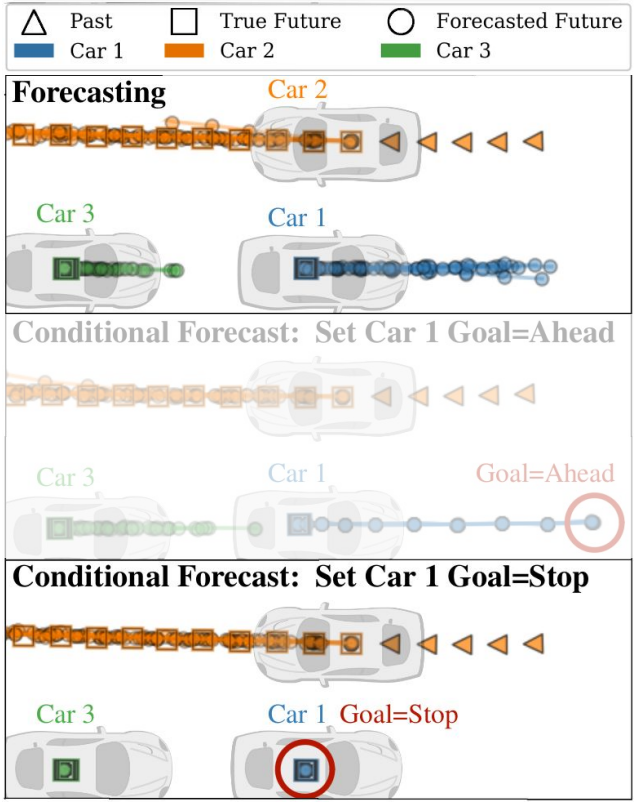


cause

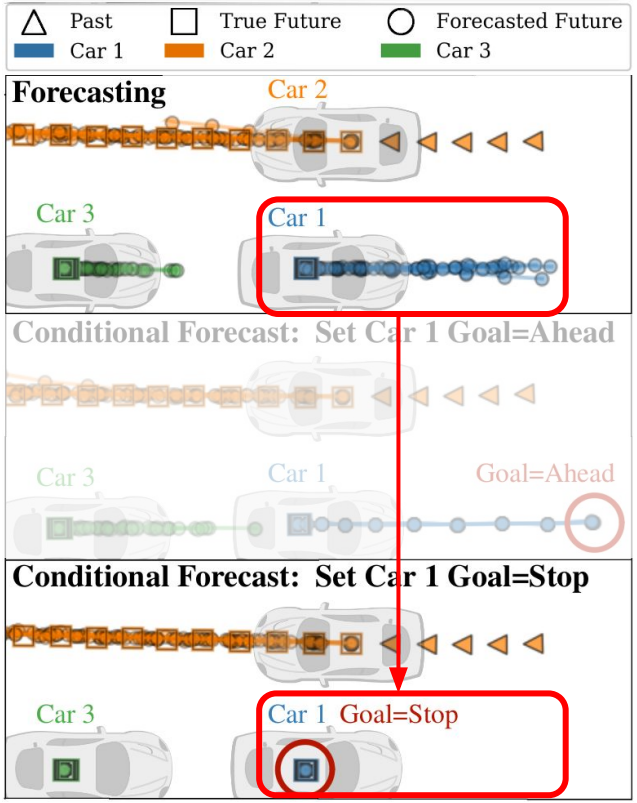
Goal-Conditioned Multi-Agent Forecasting



Goal-Conditioned Multi-Agent Forecasting

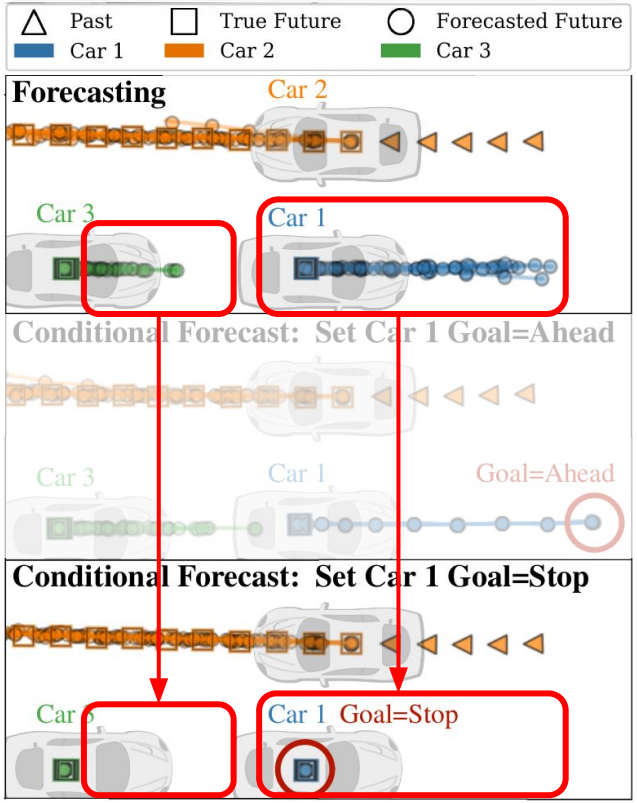


Goal-Conditioned Multi-Agent Forecasting



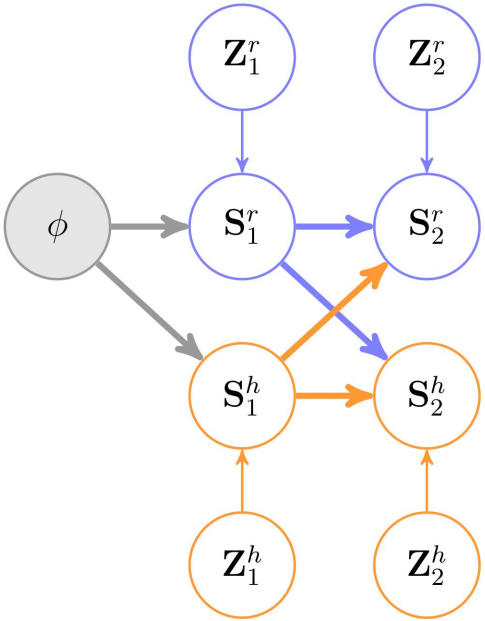
cause

Goal-Conditioned Multi-Agent Forecasting

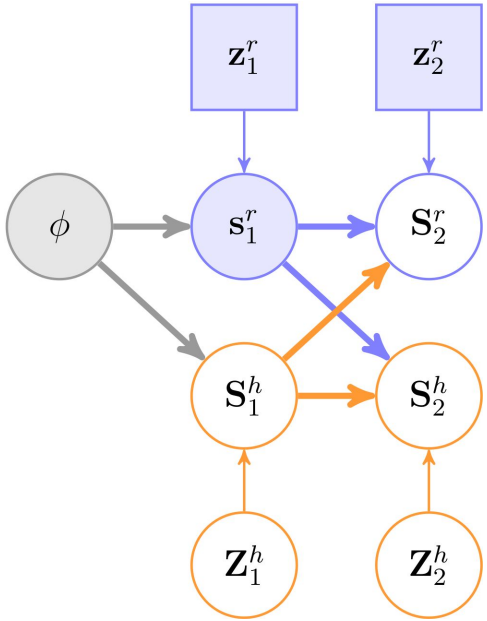


Goal-Conditioned Multi-Agent Forecasting

Forecasting

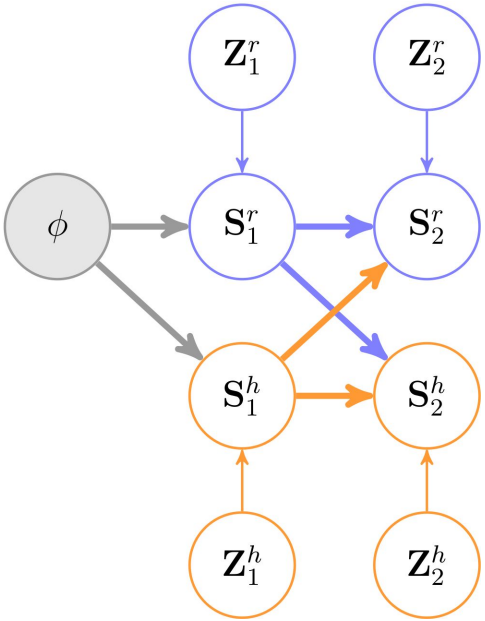


Conditional forecasting

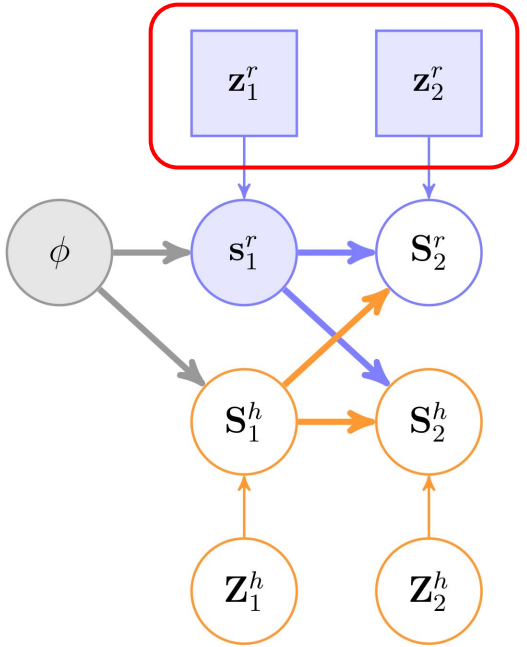


Goal-Conditioned Multi-Agent Forecasting

Forecasting



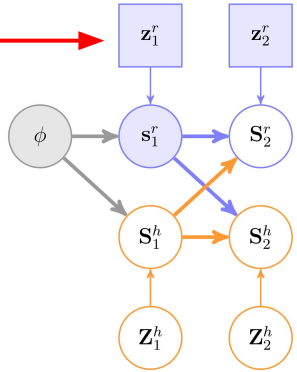
Conditional forecasting



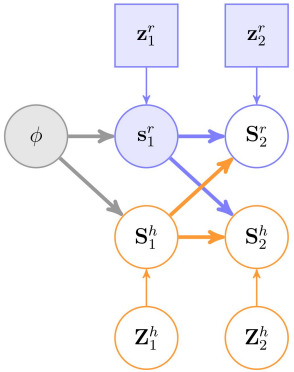
How to decide?

Goal-Conditioned Multi-Agent Forecasting

How to decide ?

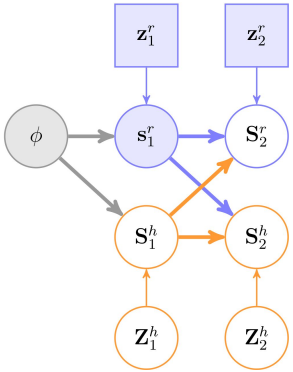


Goal-Conditioned Multi-Agent Forecasting



$$p\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{goal}, \textbf{context}\right)$$

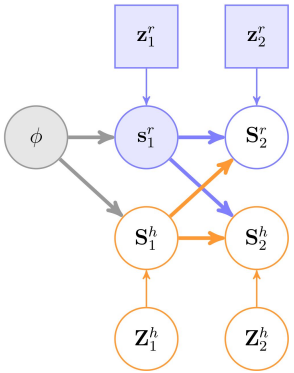
Goal-Conditioned Multi-Agent Forecasting



$$\text{states}, \text{states} = f\left(\text{latents}, \text{latents}\,;\, \text{context}\right)$$

$$p\left(\text{states}, \text{states} \,\middle|\, \text{goal}, \text{context}\right)$$

Goal-Conditioned Multi-Agent Forecasting

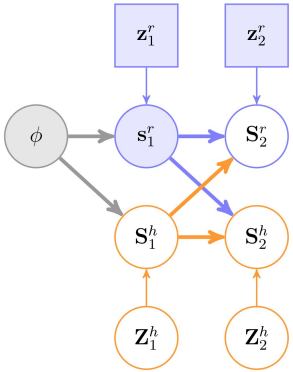


$$\textbf{states}, \textbf{states} \;=\; f\left(\textbf{latents}, \textbf{latents}\,;\, \textbf{context}\right)$$

numerics

$$\log p\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{goal}, \textbf{context}\right)$$

Goal-Conditioned Multi-Agent Forecasting



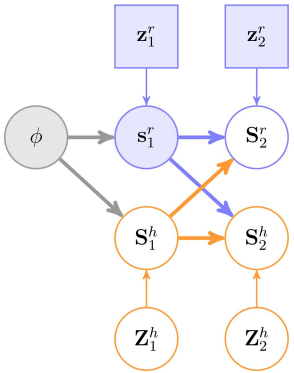
$$\mathbf{states}, \mathbf{states} = f\left(\mathbf{latents}, \mathbf{latents}\,;\, \mathbf{context}\right)$$

numerics

$$\mathbb{E}_{\mathbf{latents}}\left[\log p\left(\mathbf{states}, \mathbf{states} \,\middle|\, \mathbf{goal}, \mathbf{context}\right)\right]$$

cannot control

Goal-Conditioned Multi-Agent Forecasting



$$\text{states}, \text{states} = f\left(\text{latents}, \text{latents}\,;\, \text{context}\right)$$

numerics

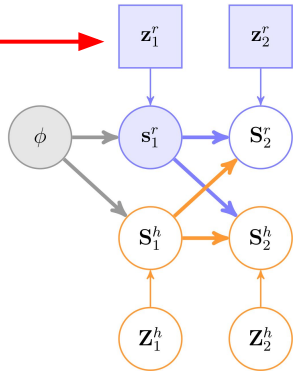
$$\arg\max_{\text{latents}} \mathbb{E}_{\text{latents}}\left[\log p\left(\text{states}, \text{states} \,\middle|\, \text{goal}, \text{context}\right)\right]$$

cannot control

can control

Goal-Conditioned Multi-Agent Forecasting

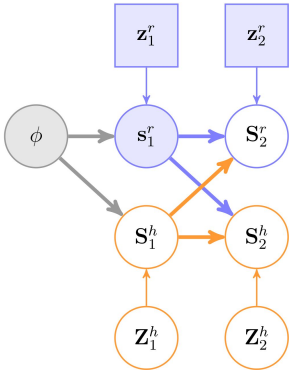
How we decide



$$\text{states}, \text{states} = f\left(\text{latents}, \text{latents}\,;\, \text{context}\right)$$

$$\text{latents}^* = \arg\max_{\text{latents}} \mathbb{E}_{\text{latents}}\left[\log p\left(\text{states}, \text{states} \,\middle|\, \text{goal}, \text{context}\right)\right]$$

Goal-Conditioned Multi-Agent Forecasting

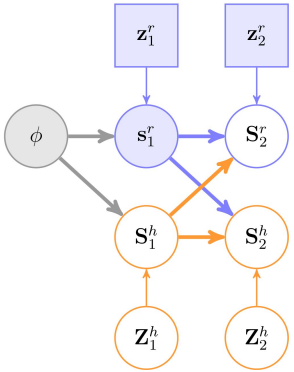


$$\textbf{states}, \textbf{states} = f\left(\textbf{latents}, \textbf{latents}\,;\, \textbf{context}\right)$$

$$\textbf{latents}^* = \arg\max_{\textbf{latents}} \mathbb{E}_{\textbf{latents}}\left[\log p\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{goal}, \textbf{context}\right)\right]$$

$$= \arg\max_{\textbf{latents}} \mathbb{E}_{\textbf{latents}}\left[\log q\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{context}\right) + \log p(\textbf{goal} \,\middle|\, \textbf{states}, \textbf{states}, \textbf{context}) - \log p\left(\textbf{goal} \,\middle|\, \textbf{context}\right)\right]$$

Goal-Conditioned Multi-Agent Forecasting



$$\textbf{states}, \textbf{states} = f\left(\textbf{latents}, \textbf{latents}\,;\, \textbf{context}\right)$$

$$\textbf{latents}^* = \arg\max_{\textbf{latents}} \mathbb{E}_{\textbf{latents}}\left[\log p\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{goal}, \textbf{context}\right)\right]$$

$$= \arg\max_{\textbf{latents}} \mathbb{E}_{\textbf{latents}}\left[\log q\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{context}\right) + \log p(\textbf{goal} \,|\, \textbf{states}, \textbf{states}, \textbf{context}) - \log p\left(\textbf{goal} \,\middle|\, \textbf{context}\right)\right]$$

$$= \arg\max_{\textbf{latents}} \mathbb{E}_{\textbf{latents}}\Big[\underbrace{\log q\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{context}\right)}_{\text{multi-agent prior}} + \underbrace{\log p\left(\textbf{goal} \,\middle|\, \textbf{states}, \textbf{states}, \textbf{context}\right)}_{\text{goal likelihood}}\Big]$$

Future work: how to respond to “out-of-distribution” scenes?

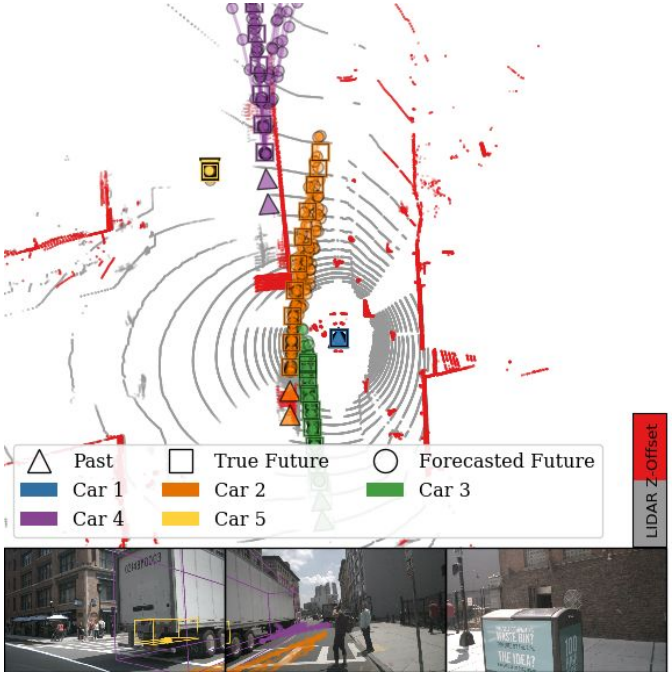
Thank you!

Single-agent forecasting + control
Deep Imitative Models for Flexible Inference, Planning, and Control
Nicholas Rhinehart, Rowan McAllister, Sergey Levine
http://imitate.ml

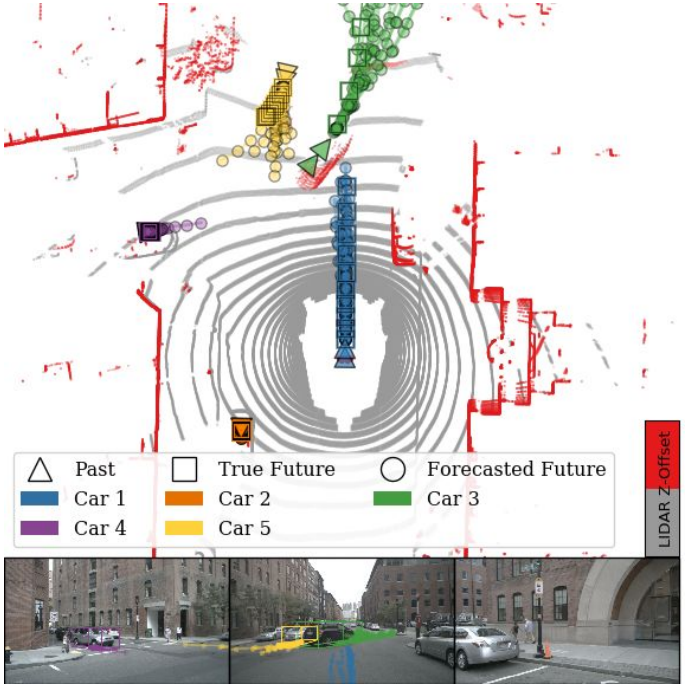
Multi-agent forecasting
PRECOG: PREdiction Conditioned On Goals in Visual Multi-Agent Settings
Nicholas Rhinehart, Rowan McAllister, Kris Kitani, Sergey Levine
http://precog.ml

Bayesian single-agent forecasting + control
(in progress)

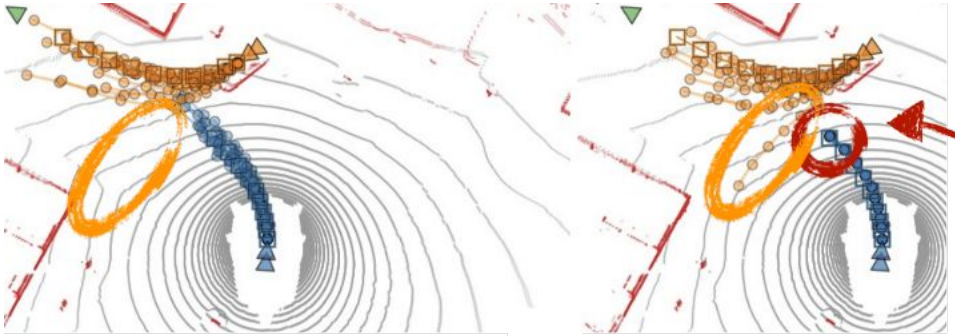
Forecasting with nuScenes data (Singapore + Boston)



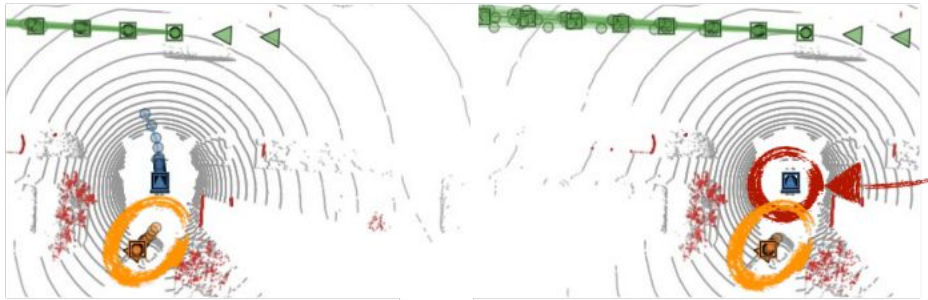
Car 2 is predicted to overtake **Car 1**, which itself is forecasted to continue to wait for pedestrians and **Car 2**.



Car 4 is predicted to wait for a clear intersection, and **Car 5** is predicted to either start turning or continue straight.



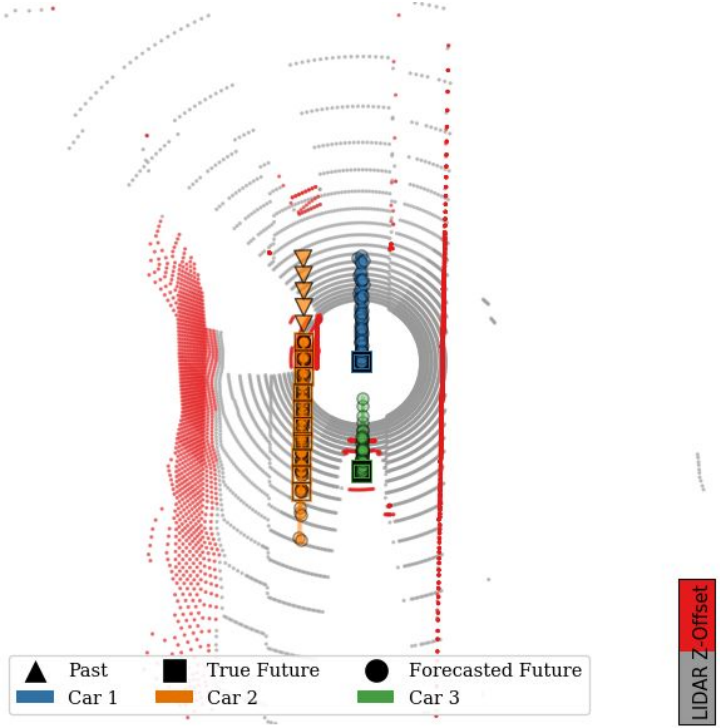
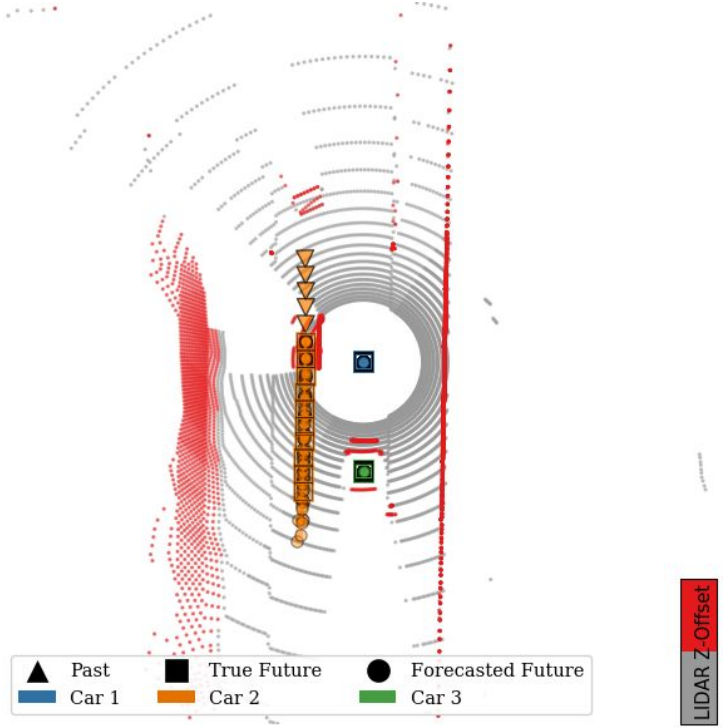
Forecasting PRECOG



Forecasting PRECOG

Goal-Conditioned Forecasting

Forecasting



Car 1 Forecasts Stopping