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 ?	?	✗	✓	

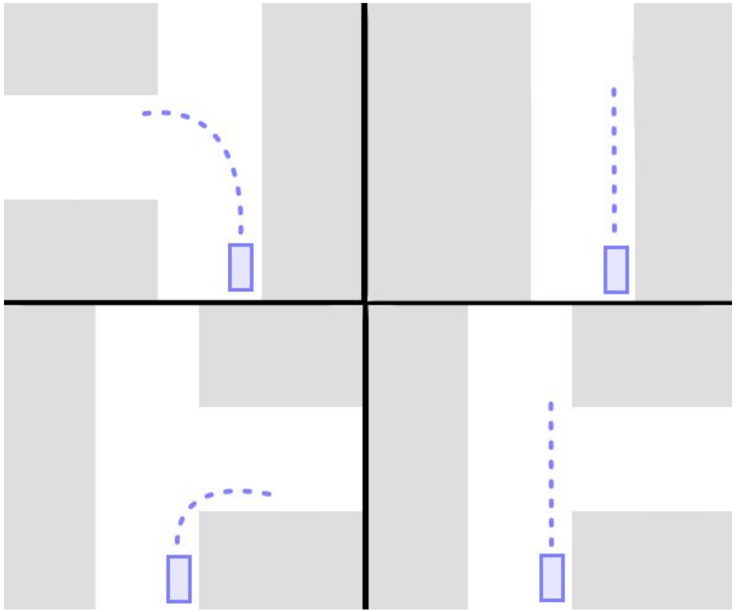
Q: Can robots **safely learn** to drive **suburban roads** in **interpretable** ways to **new goals** ?

	(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** ?

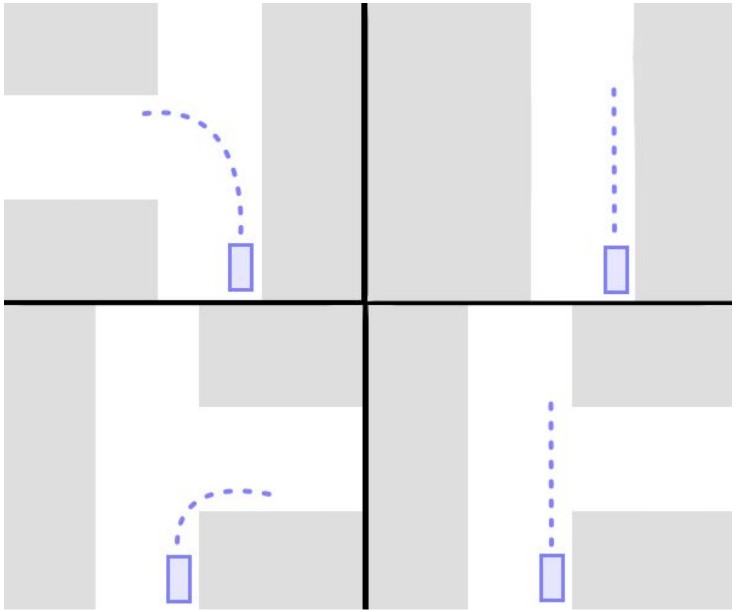
	(trains offline)	(no rewards required)	(plans)	(dyn. model)
Without learning ?	✔	✘	✔	✔
Behavior cloning ?	?	✔	✘	✘
Model-based RL ?	?	✘	✔	✔
Imitative Models (ours)	✔	✔	✔	✔

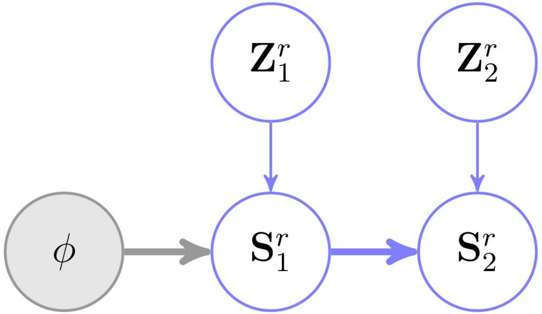
Modelling Expert Drivers



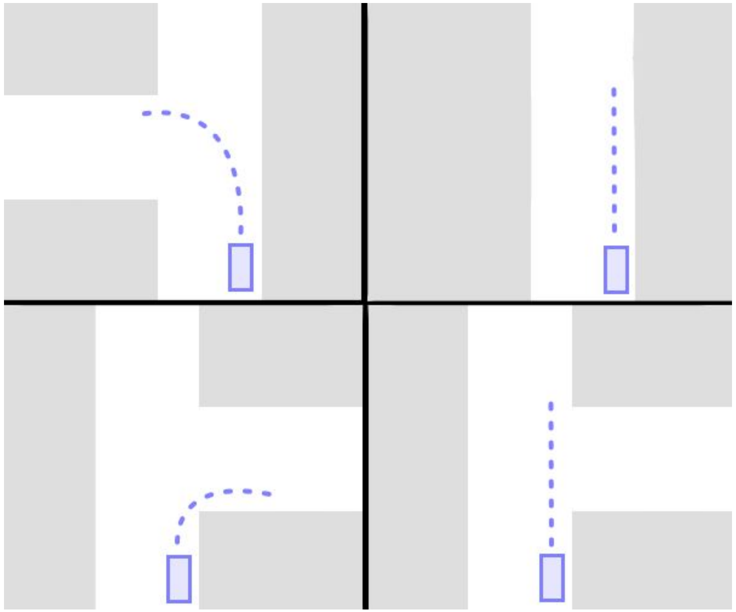
Rhinehart, Kitani, Vernaza **http://r2p2.ml**

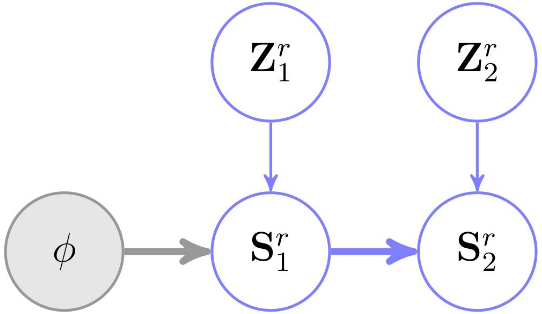
Modelling Expert Drivers





Modelling Expert Drivers

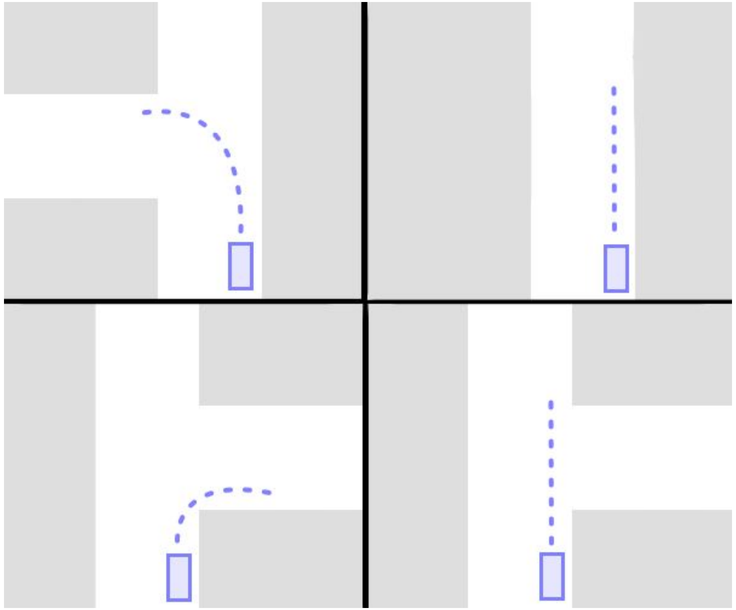


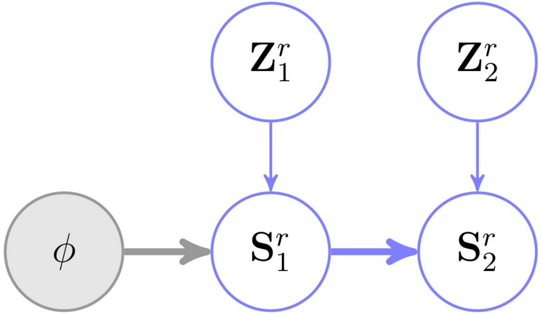


$$\textbf{states} = f\,(\textbf{latents}\,;\textbf{context})$$

Rhinehart, Kitani, Vernaza **http://r2p2.ml**

Modelling Expert Drivers

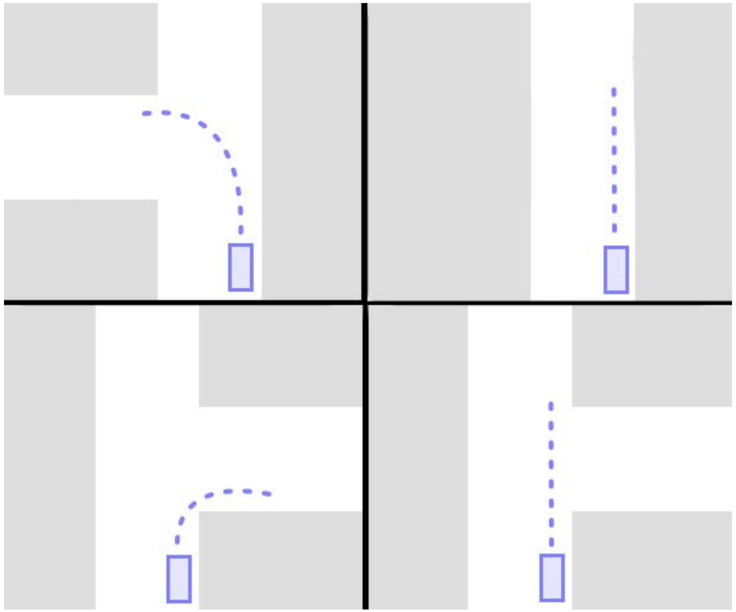


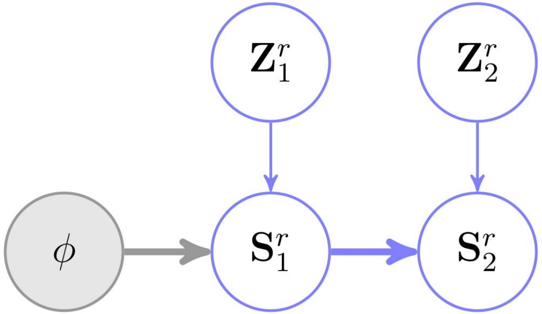


$$\begin{aligned} \textbf{states} &= f\left(\textbf{latents}\,;\textbf{context}\right) \\ \textbf{latents} &= f^{-1}\left(\textbf{states}\,;\textbf{context}\right) \end{aligned}$$

Rhinehart, Kitani, Vernaza **http://r2p2.ml**

Modelling Expert Drivers



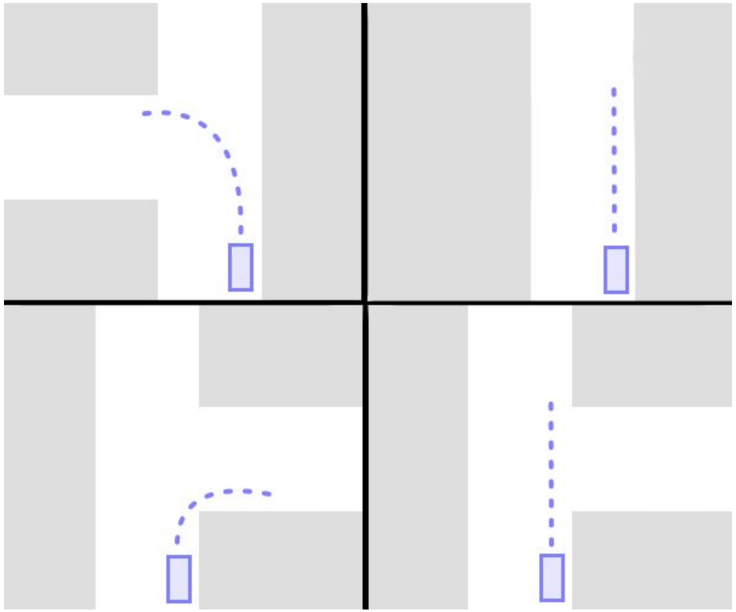


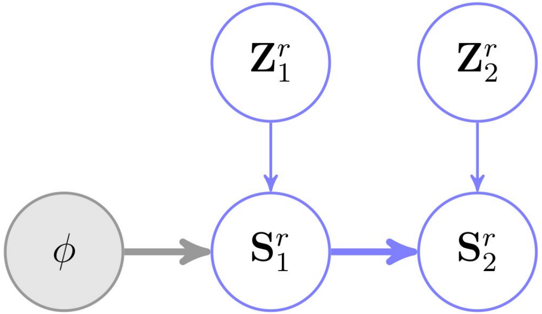
$$\begin{aligned}\textbf{states} &= f\left(\textbf{latents}\,;\textbf{context}\right)\\ \textbf{latents} &= f^{-1}\left(\textbf{states}\,;\textbf{context}\right)\end{aligned}$$

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

Rhinehart, Kitani, Vernaza **http://r2p2.ml**

Modelling Expert Drivers





$$\begin{aligned} \textbf{states} &= f\,(\textbf{latents}\,;\textbf{context}) \\ \textbf{latents} &= f^{-1}\,(\textbf{states}\,;\textbf{context}) \end{aligned}$$

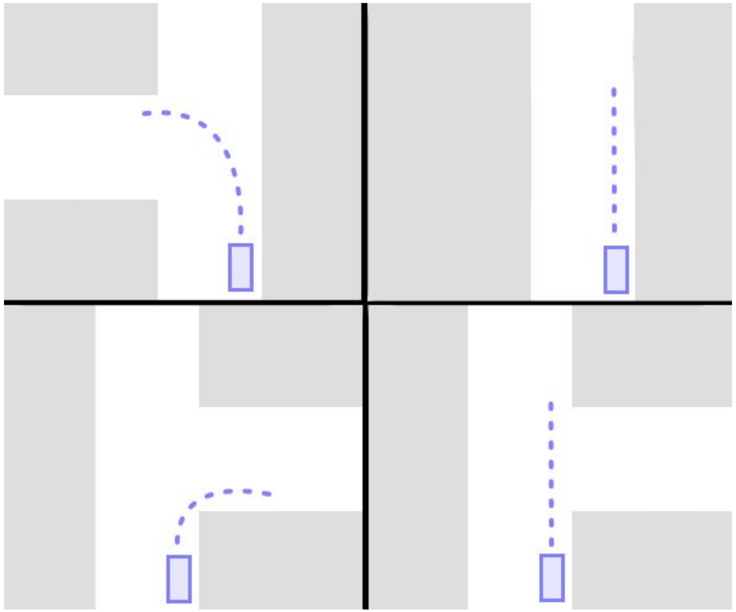
$$q\,(\textbf{states}\,|\,\textbf{context}) = \frac{\mathcal{N}\,(\textbf{latents}; 0, I)}{\left|\det \frac{\partial f}{\partial\,\textbf{latents}}\right|}$$

Can **plan** in this distribution!

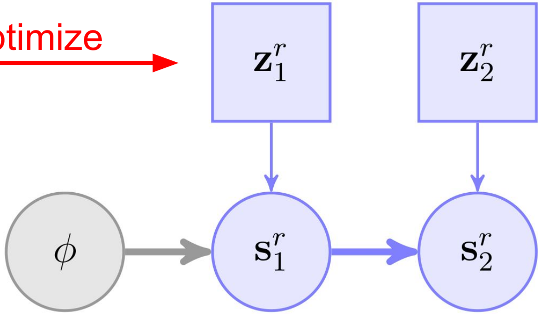
new idea

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

Modelling Expert Drivers



optimize



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

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

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

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

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

Planning “Expert-like” Motions to our own Goals

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

Drive safely...

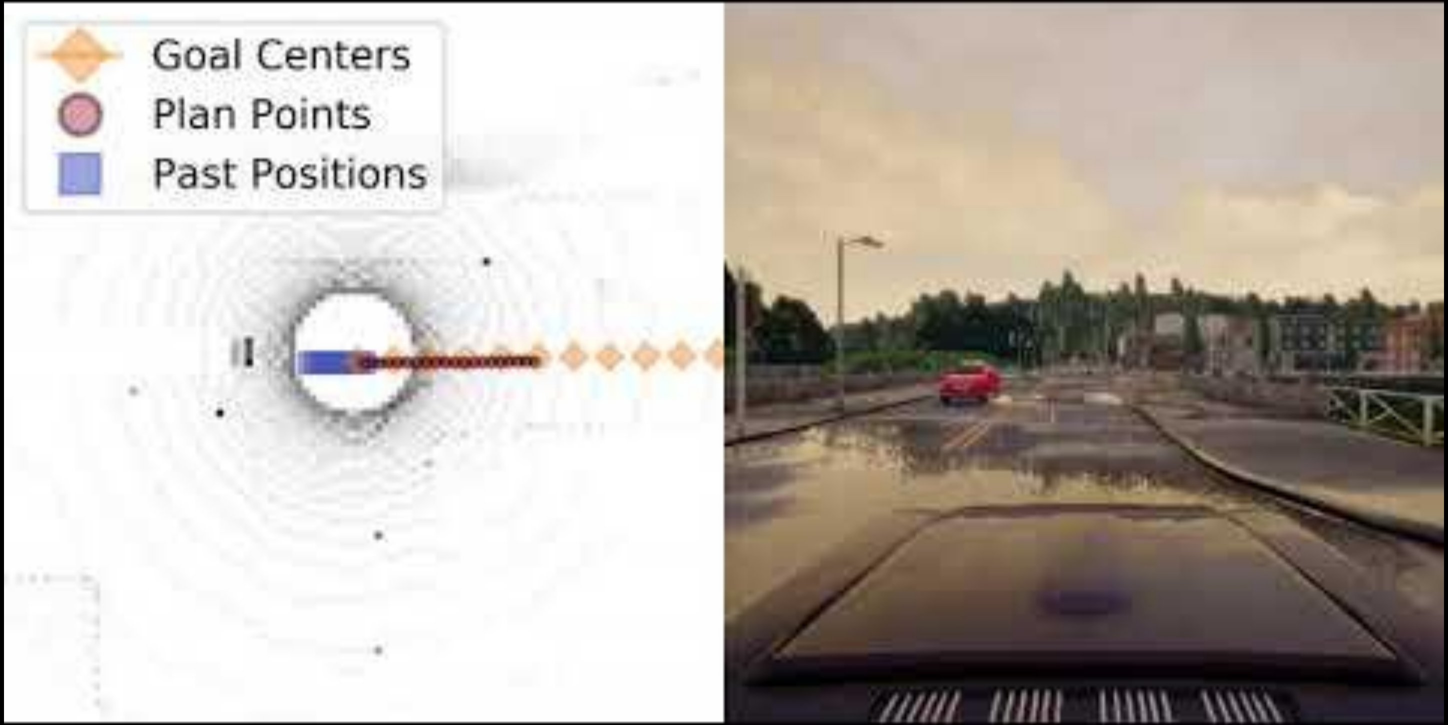
Planning “Expert-like” Motions to our own Goals

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

Drive safely...

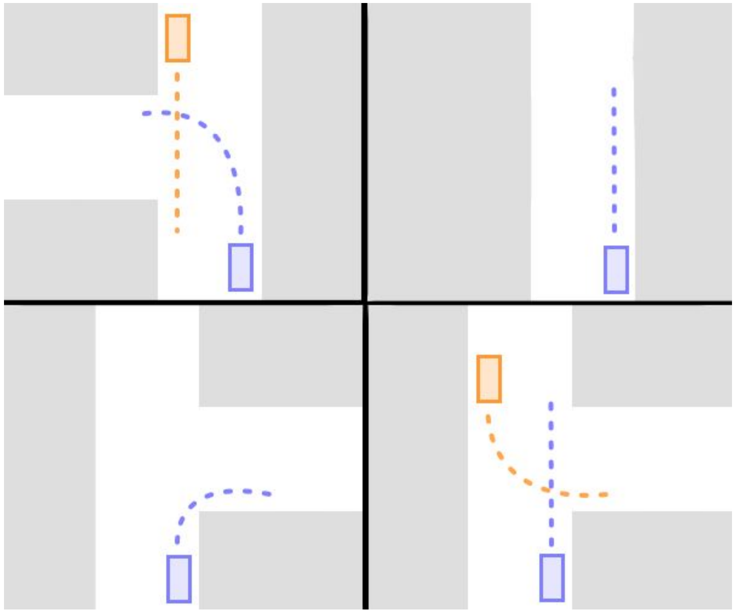
...and get to our destination!

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

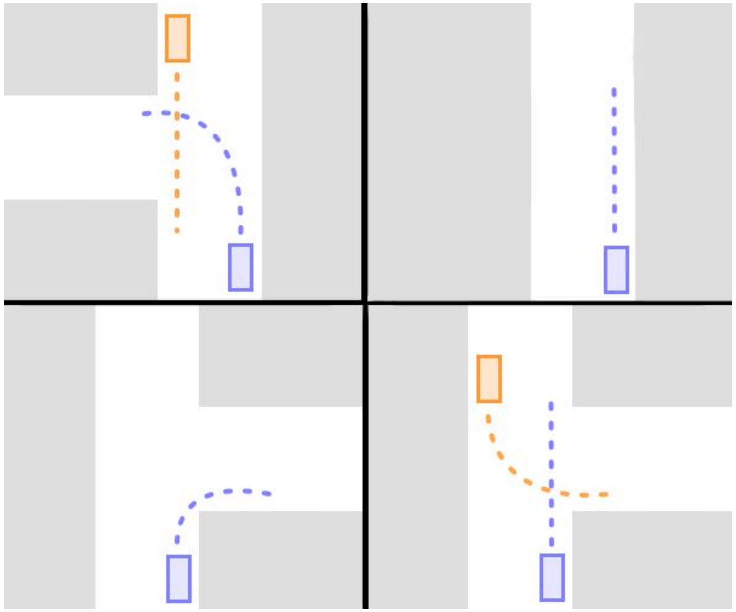

Goal Centers
Plan Points
Past Positions

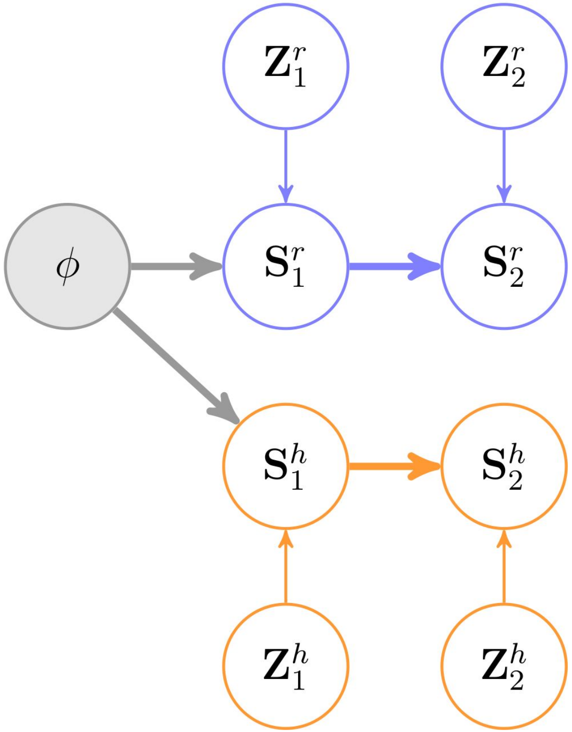
Goal Centers
Plan Points
Past Positions

Modelling Multiple Expert Drivers

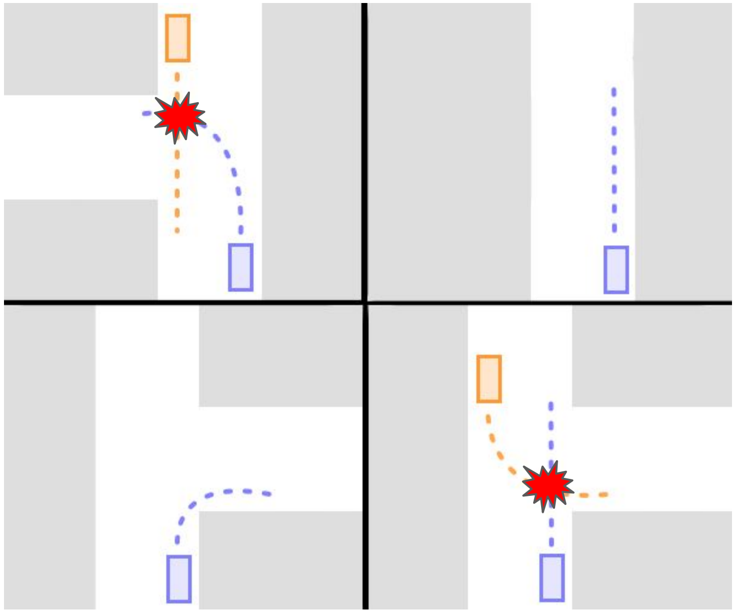


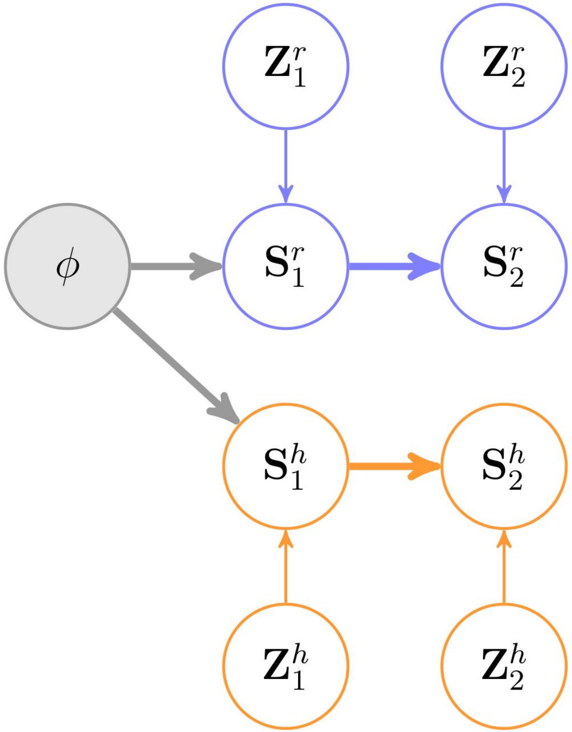
Modelling Multiple Expert Drivers



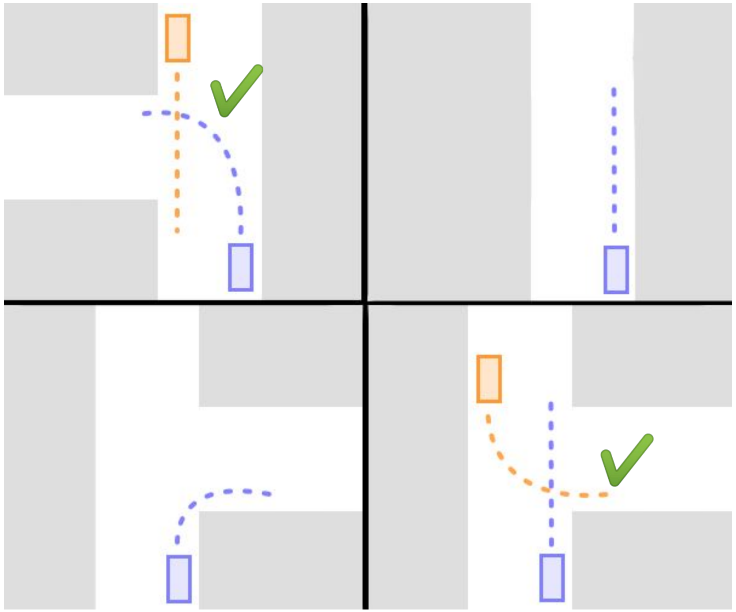


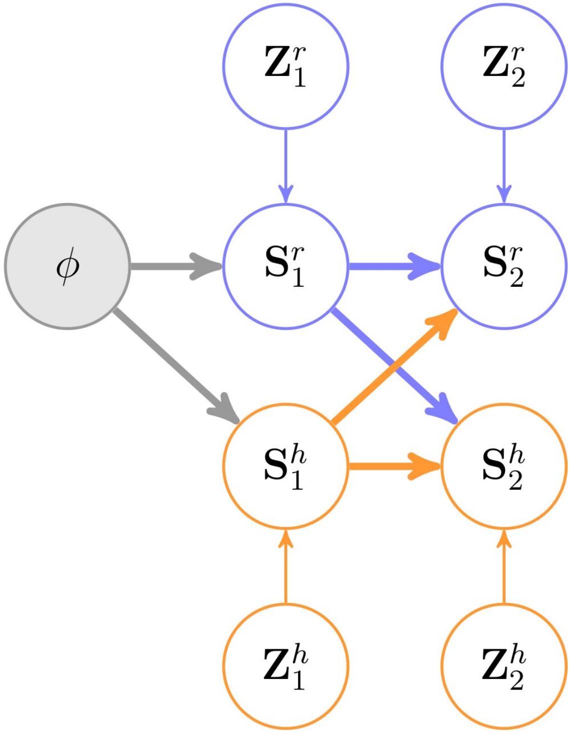
Modelling Multiple Expert Drivers





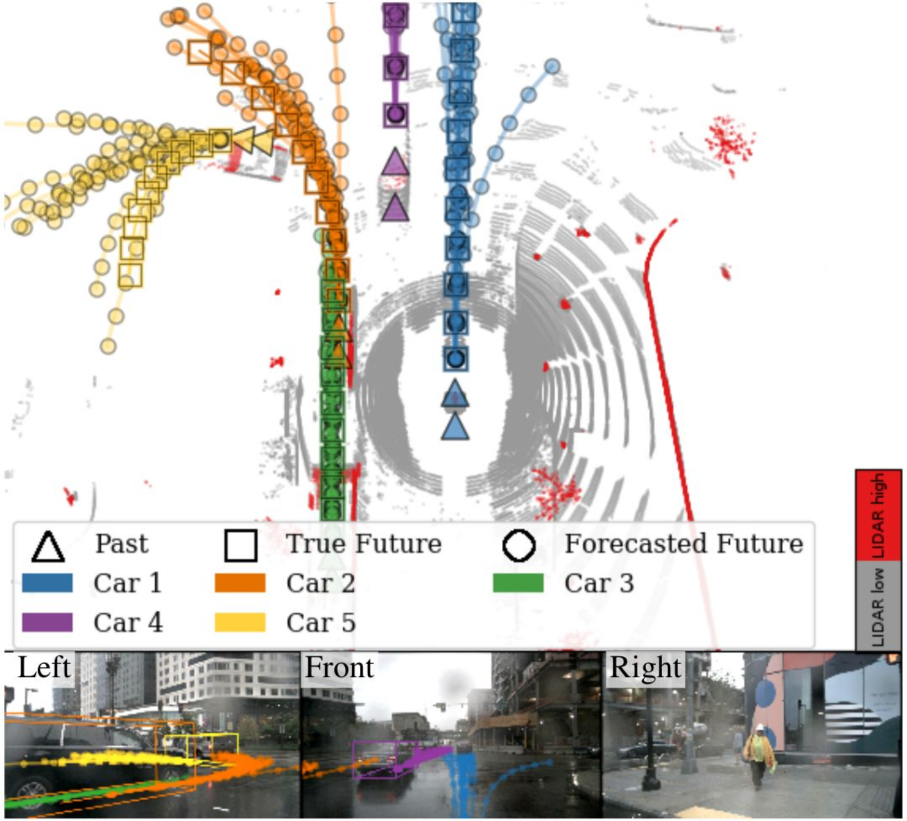
Modelling Multiple Expert Drivers



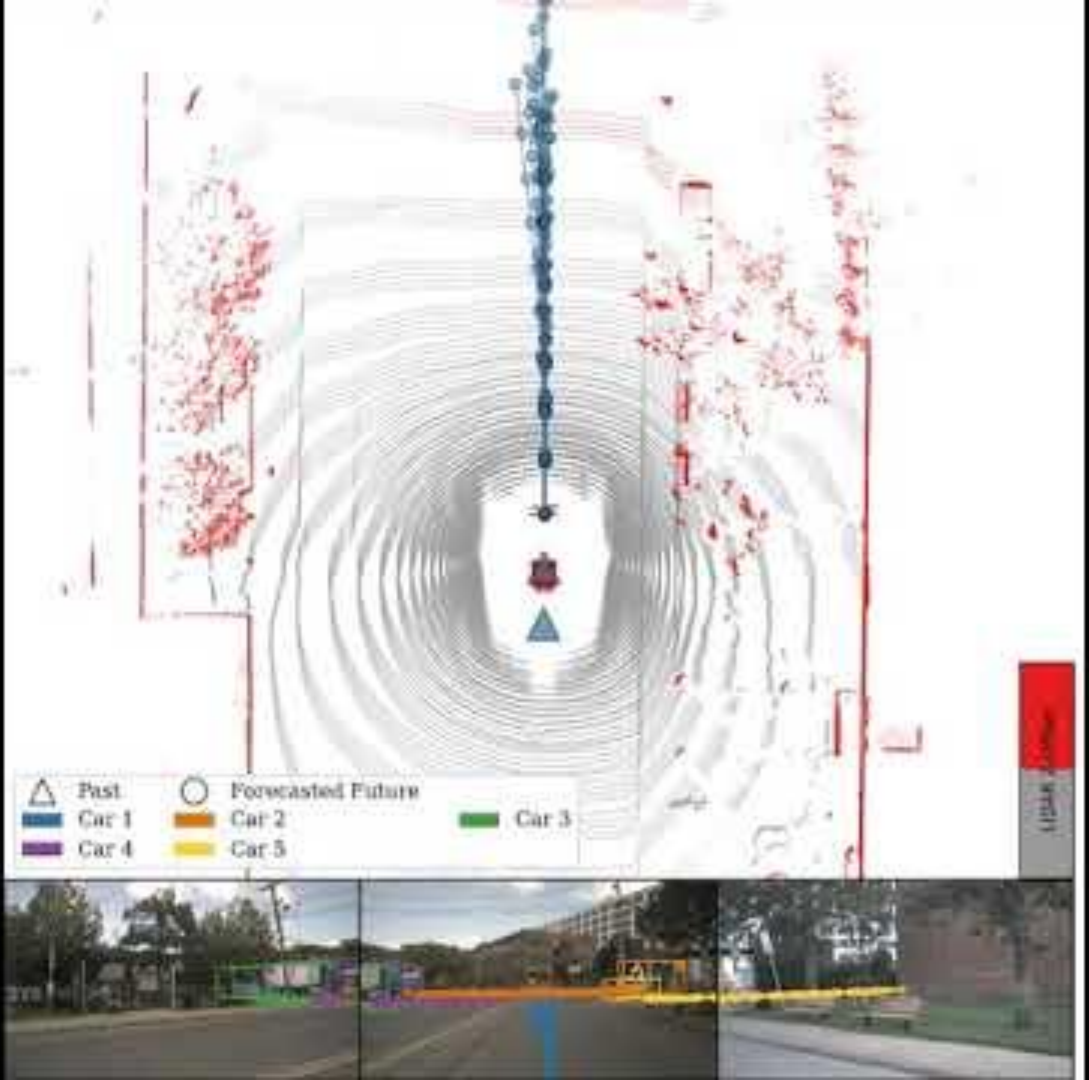


Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Forecasting with nuScenes data (Singapore + Boston)

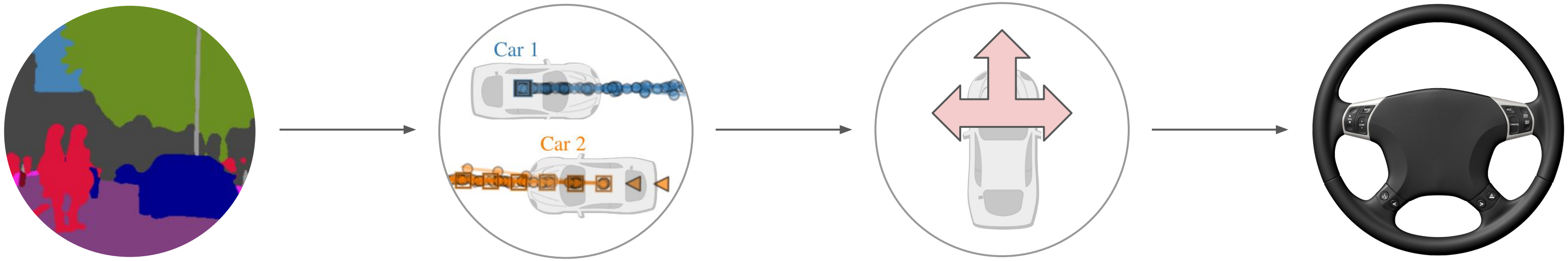


Rhinehart, McAllister, Kitani, Levine **http://precog.ml**


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

Multi-Agent Planning

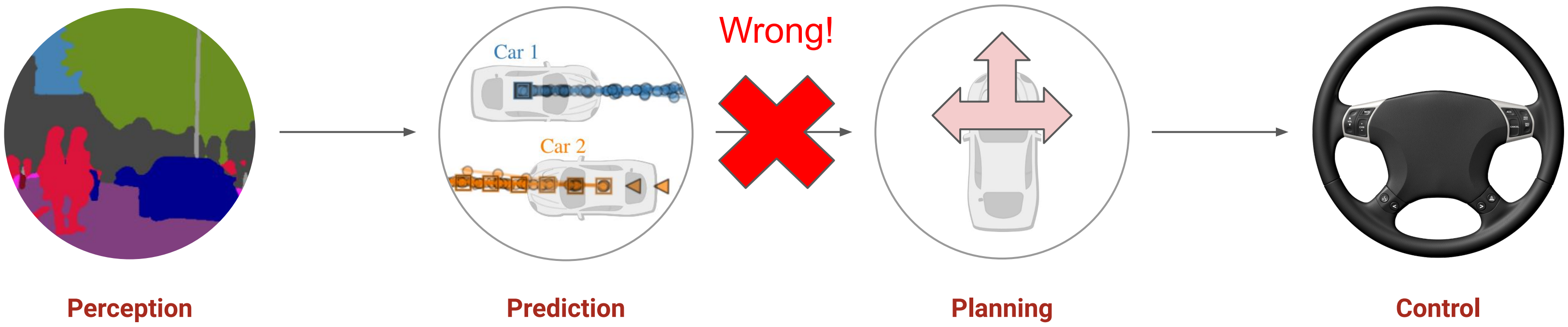
Q: but how should ***autonomous vehicles*** predict other agents...**online**?



Perception **Prediction** **Planning** **Control**

Multi-Agent Planning

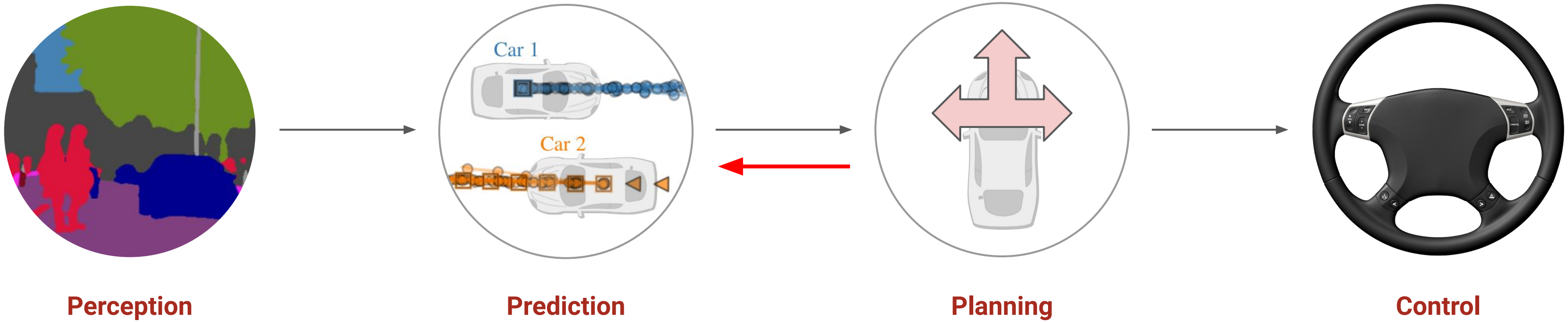
Q: but how should ***autonomous vehicles*** predict other agents...**online**?



Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Multi-Agent Planning

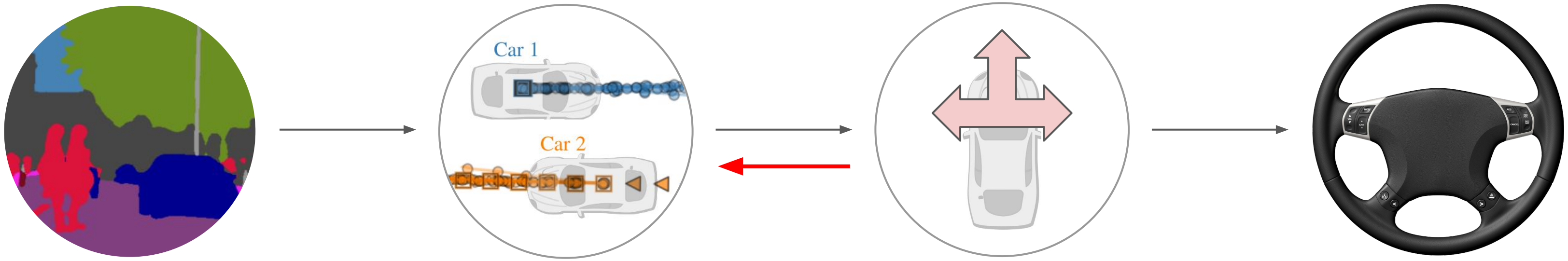
Q: but how should ***autonomous vehicles*** predict other agents...**online**?



Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Multi-Agent Planning

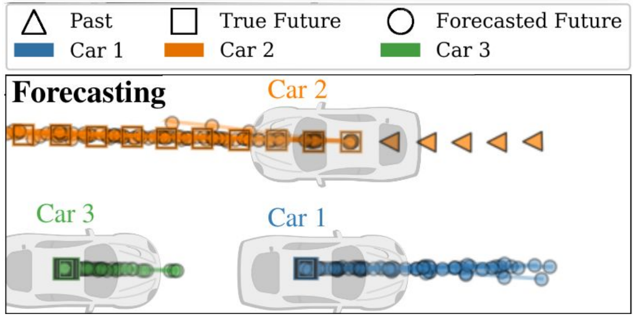
Q: but how should ***autonomous vehicles*** predict other agents...**online**?



Perception **Prediction** **Planning** **Control**

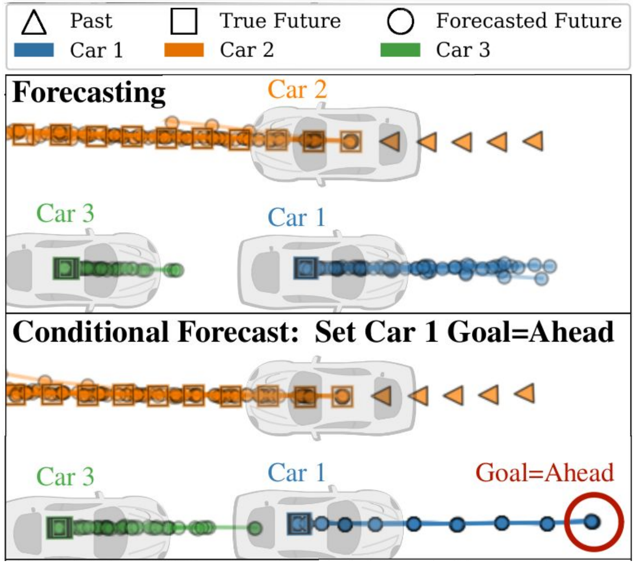
(our actions will affect others)

Goal-Conditioned Multi-Agent Forecasting



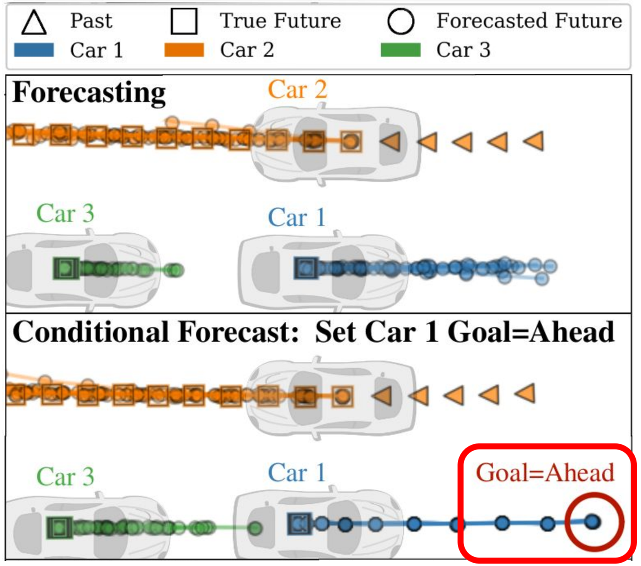
Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Goal-Conditioned Multi-Agent Forecasting



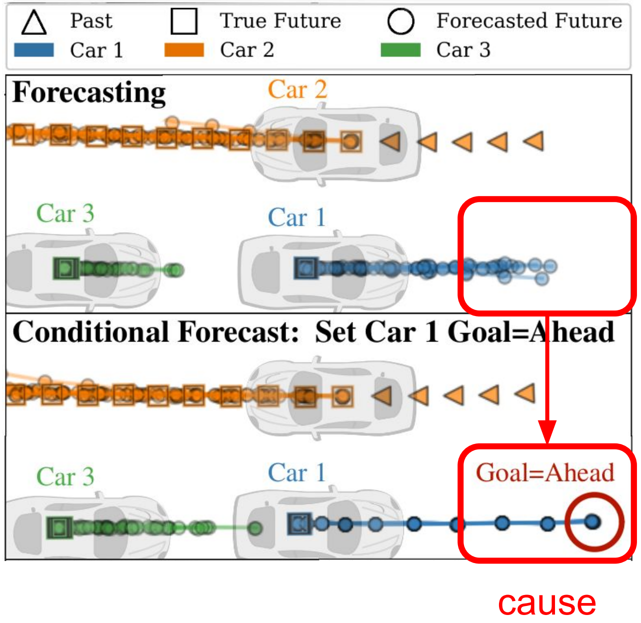
Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Goal-Conditioned Multi-Agent Forecasting



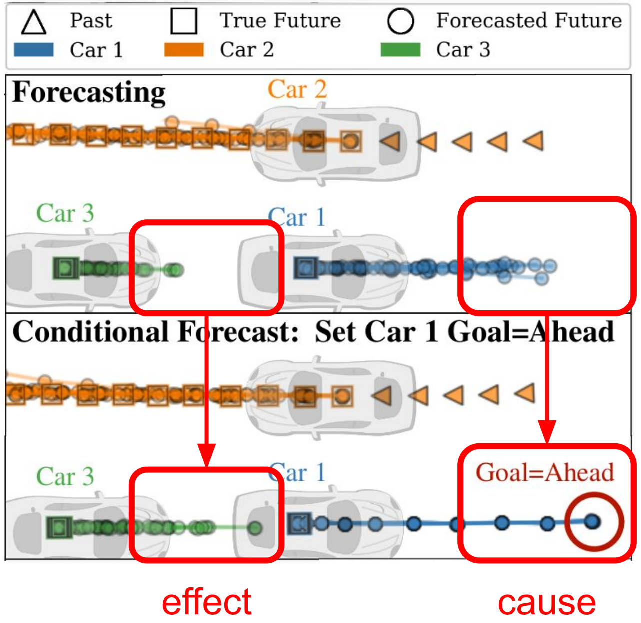
Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Goal-Conditioned Multi-Agent Forecasting

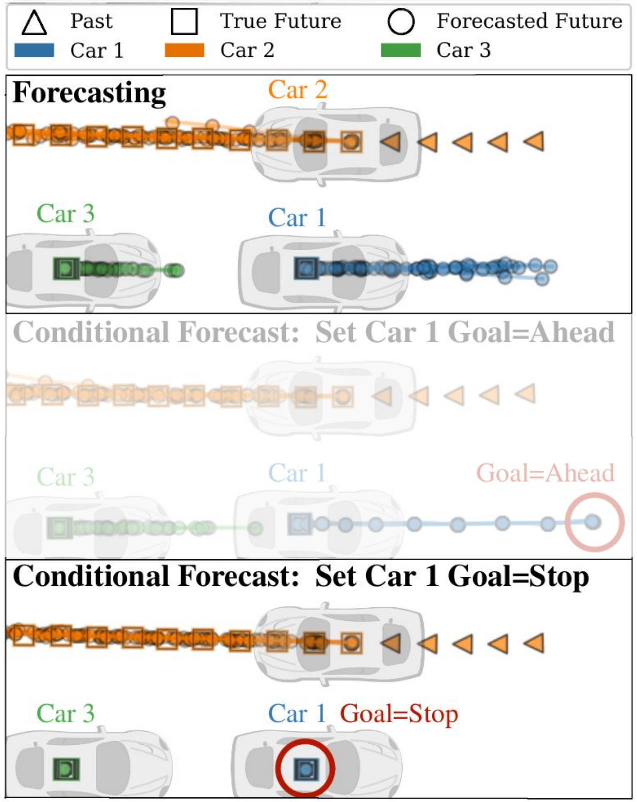


Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Goal-Conditioned Multi-Agent Forecasting

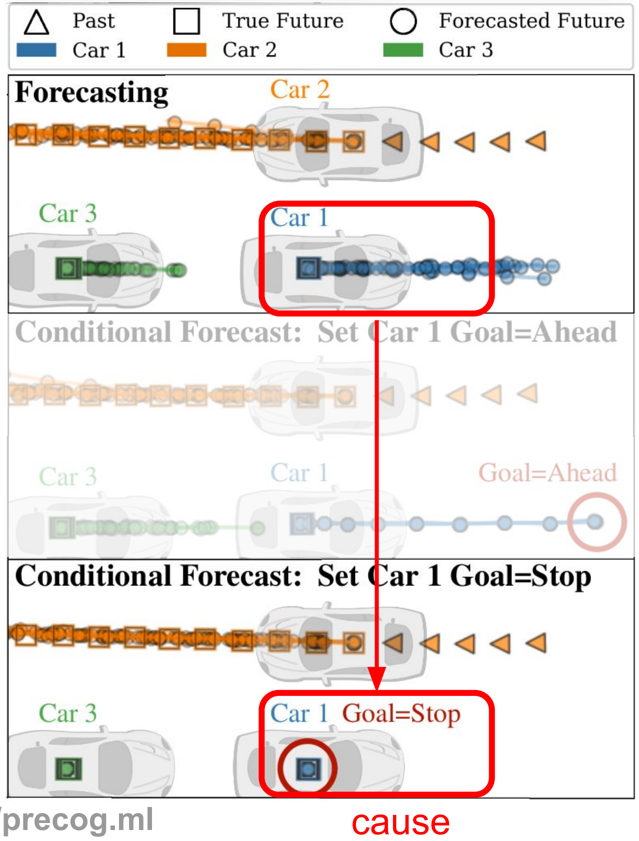


Goal-Conditioned Multi-Agent Forecasting



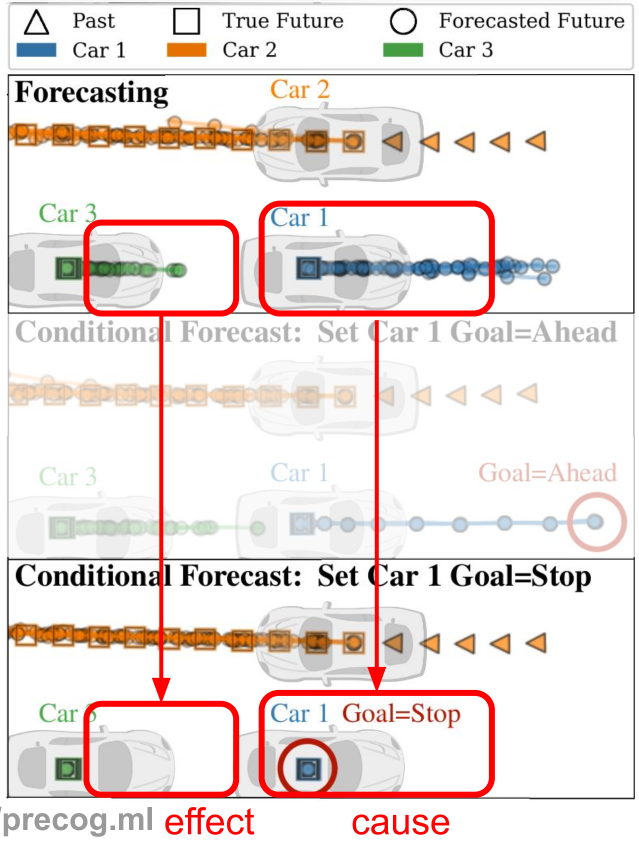
Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

Goal-Conditioned Multi-Agent Forecasting



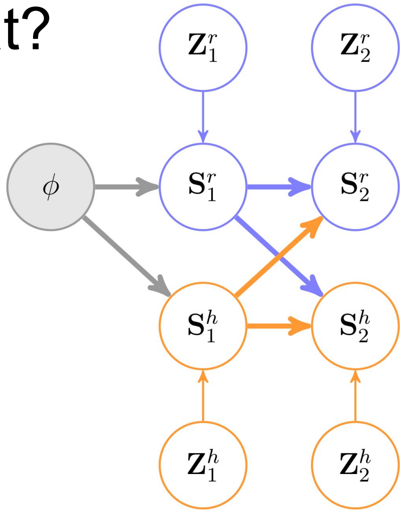
Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

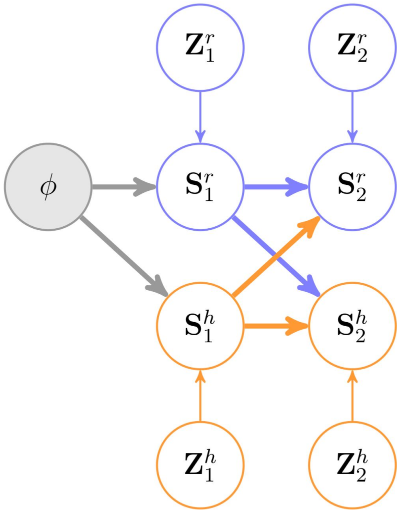
Goal-Conditioned Multi-Agent Forecasting



Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

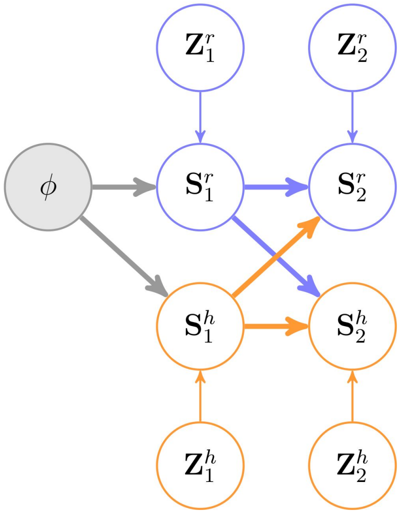
How did we do that?





Generating function:

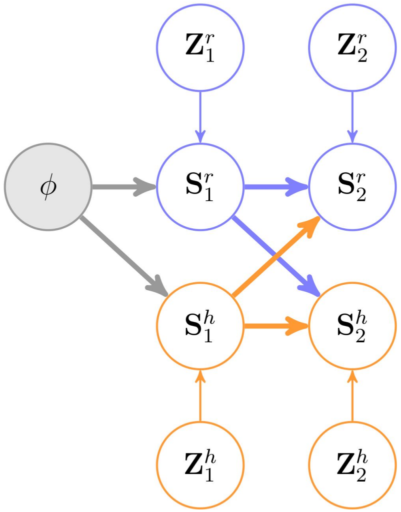
$$\mathbf{states}, \mathbf{states} \;=\; f\left(\mathbf{latents}, \mathbf{latents}\,; \mathbf{context}\right)$$



Generating function:

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

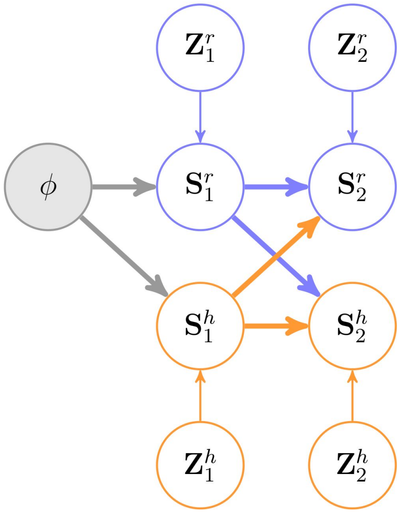
What is our Planning Objective?



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

Generating function:

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

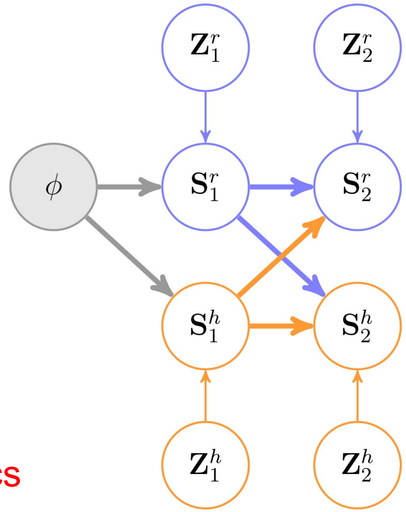


Generating function:

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

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

e.g. desired location



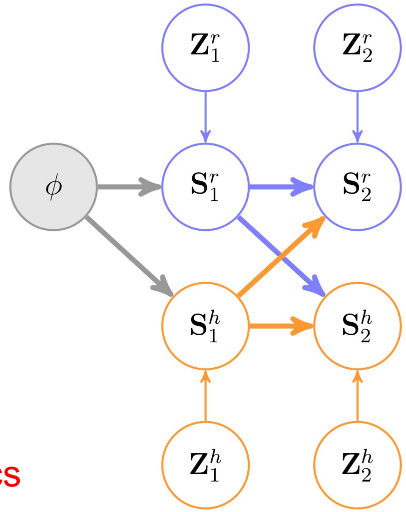
Generating function:

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

numerics

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

e.g. desired location



Generating function:

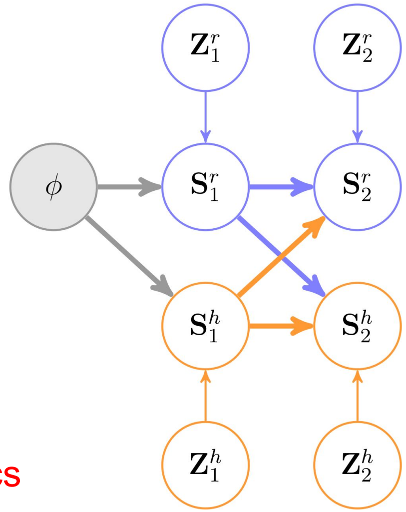
$$\textbf{states}, \textbf{states} = f\left(\textbf{latents}, \textbf{latents}\,; \textbf{context}\right)$$

numerics

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

cannot control

e.g. desired location



Generating function:

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

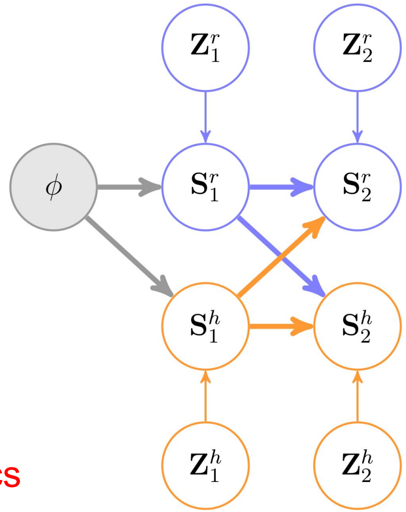
numerics

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

cannot control

e.g. desired location

can control



Generating function:

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

numerics

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

cannot control

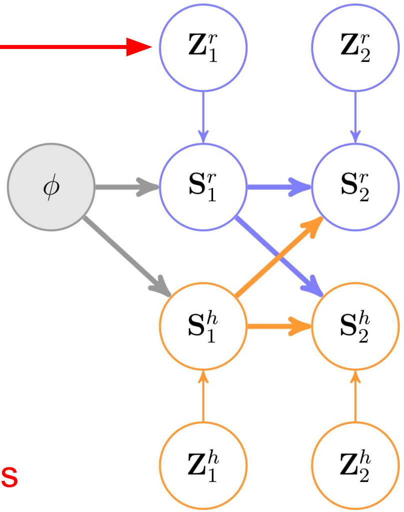
e.g. desired location

our plan

can control

Rhinehart, McAllister, Kitani, Levine **http://precog.ml**

decide



Generating function:

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

numerics

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

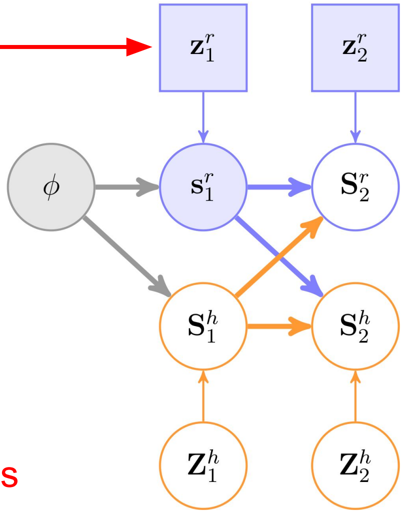
our plan

can control

cannot control

e.g. desired location

decide



Generating function:

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

numerics

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

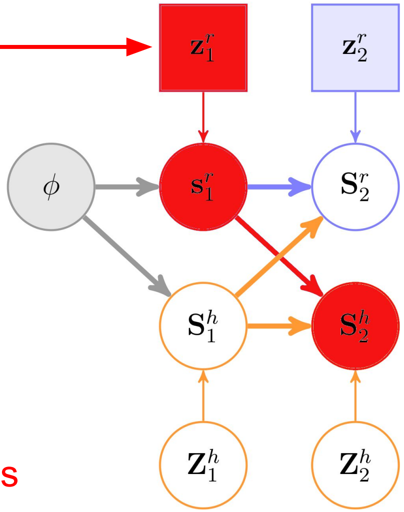
cannot control

e.g. desired location

our plan

can control

decide



cause, from robot, at t=1

effect, on human, at t=2

Generating function:

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

numerics

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

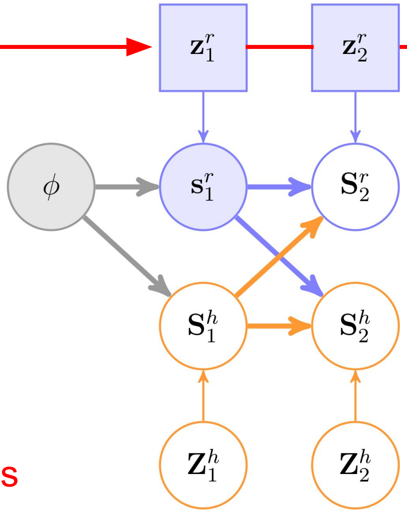
our plan

can control

cannot control

e.g. desired location

decide



Generating function:

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

~ Normal(0,1)

numerics

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

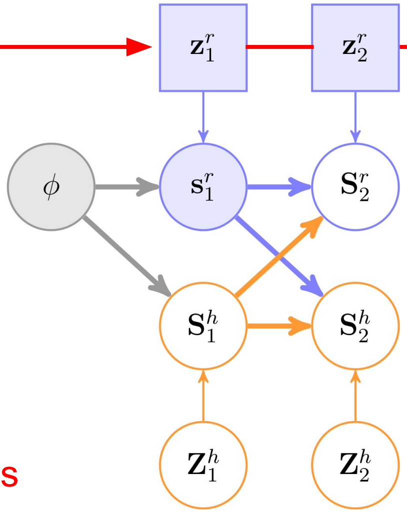
our plan

can control

cannot control

e.g. desired location

decide



Generating function:

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

~ Normal(0,1)

numerics

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

our plan

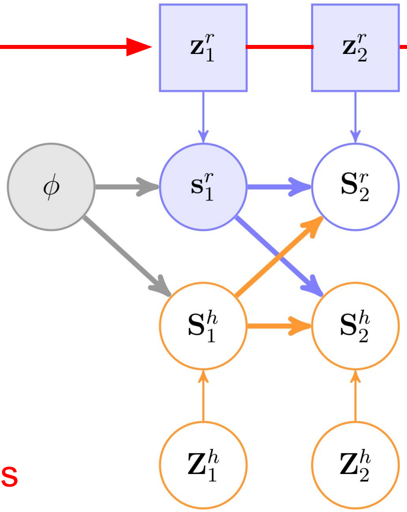
can control

cannot control

e.g. desired location

can control

decide



Generating function:

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

Predictions conditioned on **goal** !

~ Normal(0,1)

numerics

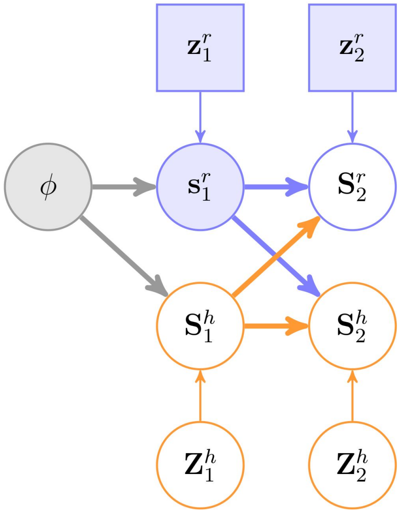
$$\boxed{\text{latents}^*} = \arg\max_{\text{latents}} \mathbb{E}_{\text{latents}}\left[\log p\left(\text{states}, \text{states} \,\middle|\, \text{goal}, \text{context}\right)\right]$$

our plan

can control

cannot control

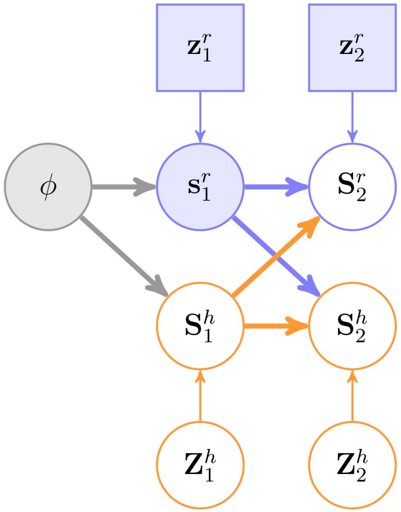
e.g. desired location



Generating function:

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

$$\begin{aligned}\textbf{latents}^* &= \underset{\textbf{latents}}{\arg\max}\, \mathbb{E}_{\textbf{latents}}\left[\log p\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{goal}, \textbf{context}\right)\right] \\ &= \underset{\textbf{latents}}{\arg\max}\, \mathbb{E}_{\textbf{latents}}\left[\log q\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{context}\right) + \log p(\textbf{goal} \,|\, \textbf{states}, \textbf{states}, \textbf{context}) - \log p\left(\textbf{goal} \,\middle|\, \textbf{context}\right)\right]\end{aligned}$$



Generating function:

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

$$\begin{aligned}
\textbf{latents}^* &= \underset{\textbf{latents}}{\arg\max}\, \mathbb{E}_{\textbf{latents}}\left[\log p\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{goal}, \textbf{context}\right)\right] \\
&= \underset{\textbf{latents}}{\arg\max}\, \mathbb{E}_{\textbf{latents}}\left[\log q\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{context}\right) + \log p(\textbf{goal} \,|\, \textbf{states}, \textbf{states}, \textbf{context}) - \log p\left(\textbf{goal} \,\middle|\, \textbf{context}\right)\right] \\
&= \underset{\textbf{latents}}{\arg\max}\, \mathbb{E}_{\textbf{latents}}\left[\underbrace{\log q\left(\textbf{states}, \textbf{states} \,\middle|\, \textbf{context}\right)}_{\text{multi-agent prior}} + \underbrace{\log p\left(\textbf{goal} \,\middle|\, \textbf{states}, \textbf{states}, \textbf{context}\right)}_{\text{goal likelihood}}\right]
\end{aligned}$$

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

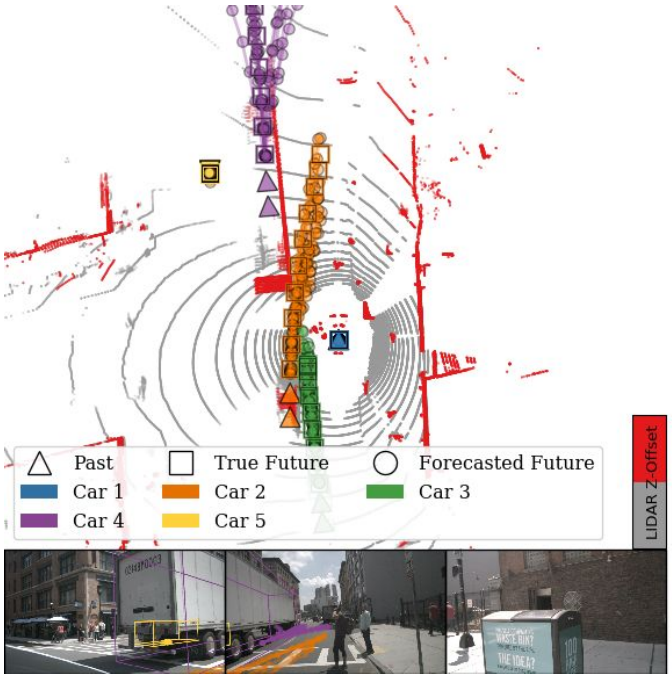
Thank you!

Single-agent forecasting + control
Deep Imitative Models for Flexible Inference, Planning, and Control
Nicholas Rhinehart, Rowan McAllister, Sergey Levine
arXiv 2018 http://imitate.ml

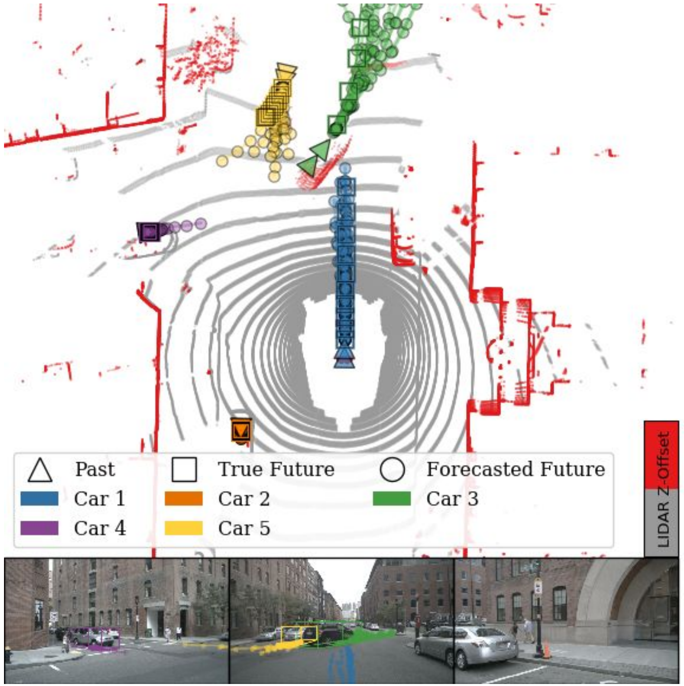
Multi-agent forecasting
PRECOG: PREdiction Conditioned On Goals in Visual Multi-Agent Settings
Nicholas Rhinehart, Rowan McAllister, Kris Kitani, Sergey Levine
ICCV 2019 http://precog.ml

Bayesian single-agent forecasting + control
(in progress)

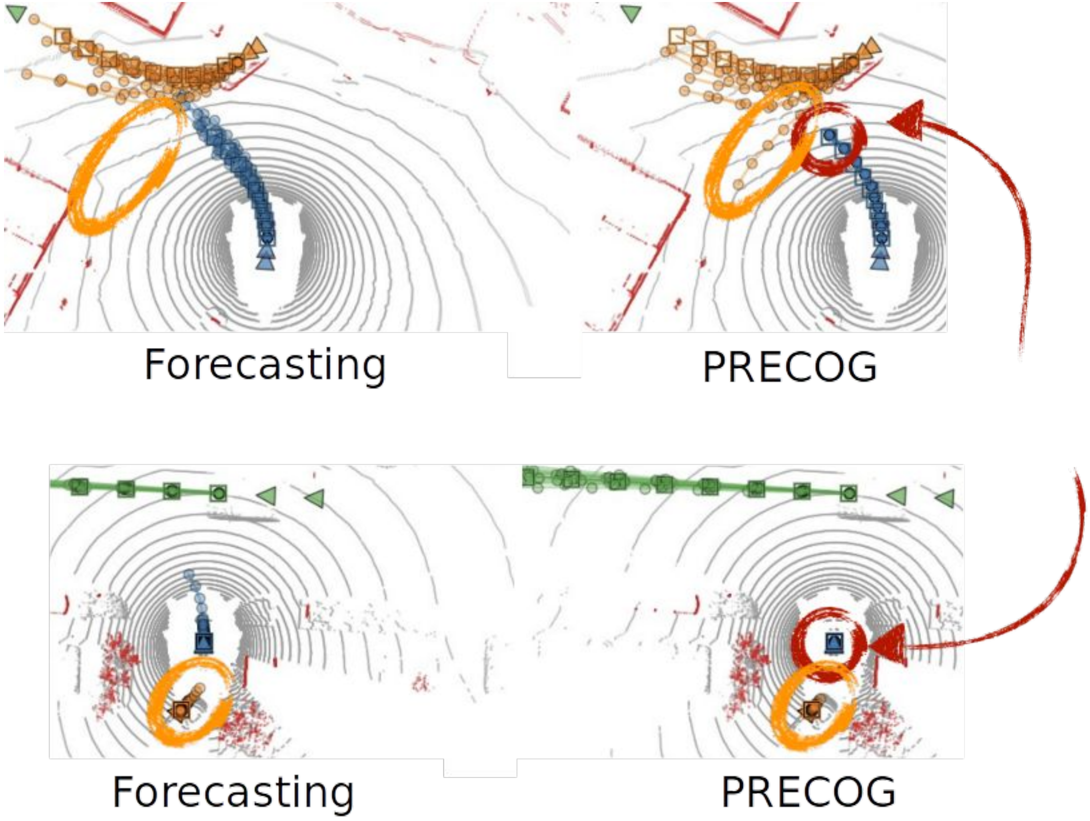
Forecasting with nuScenes data (Singapore + Boston)



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