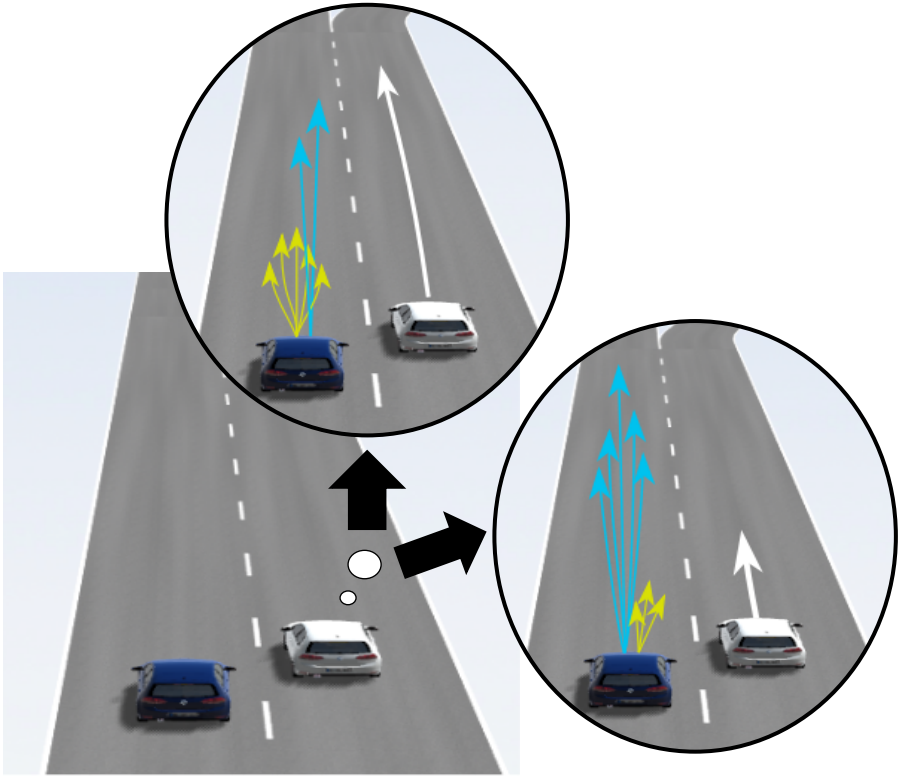
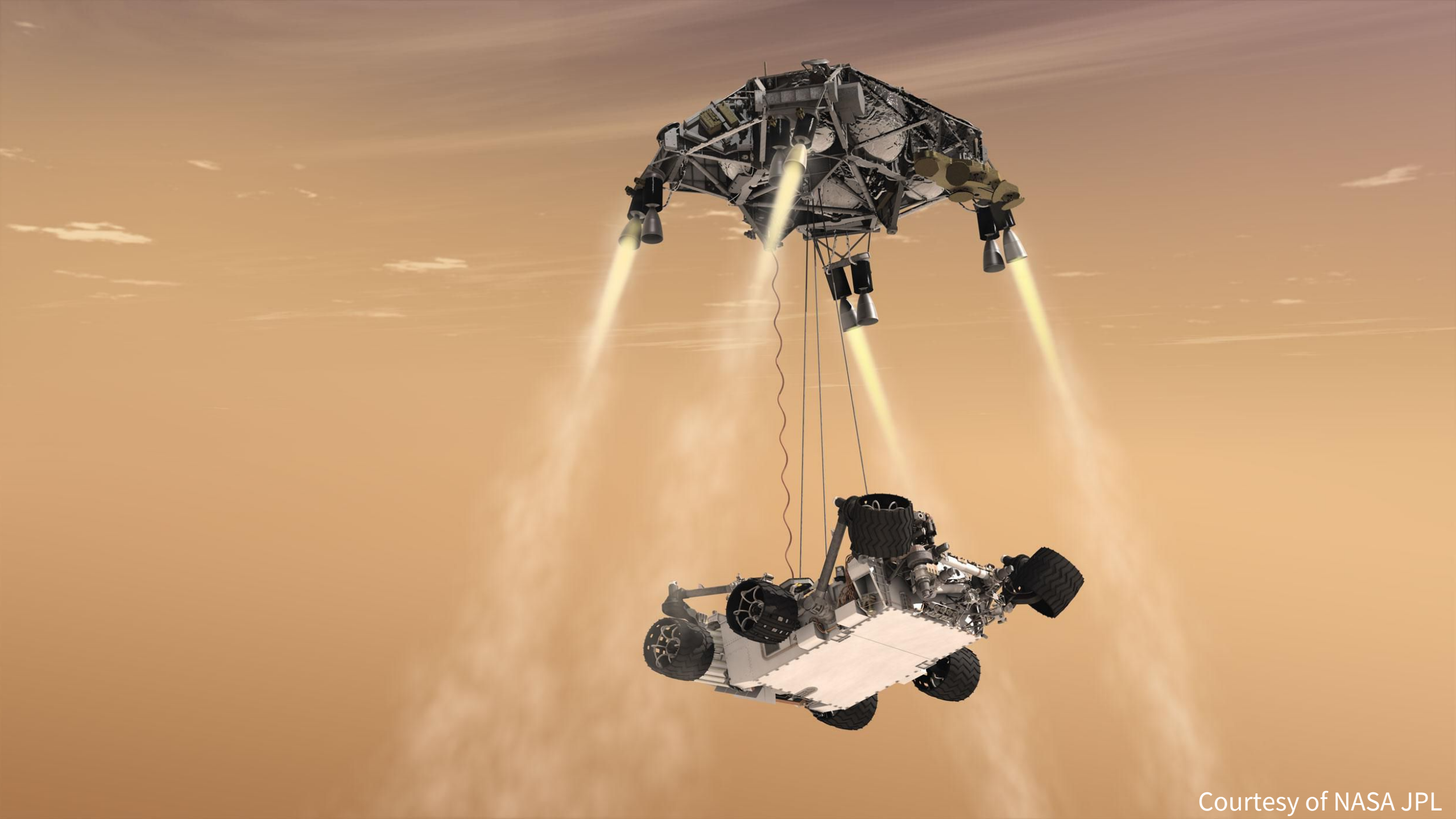
Using AI to Deal with Unpredictable Humans on the Road

Marco Pavone
Autonomous Systems Laboratory
Department of Aeronautics and Astronautics
Stanford University

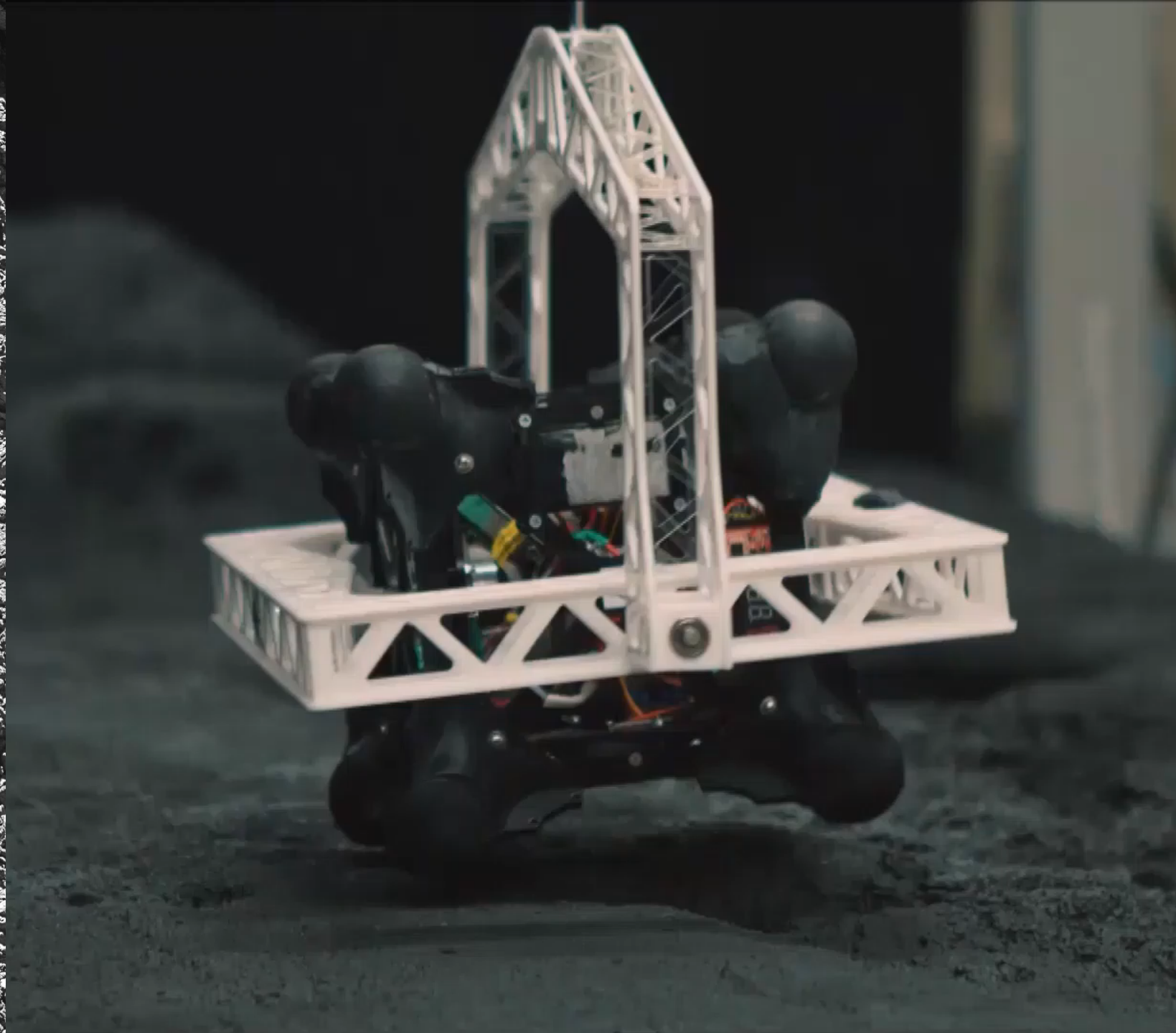
Talk@Autoware Meetup
June 24th, 2020

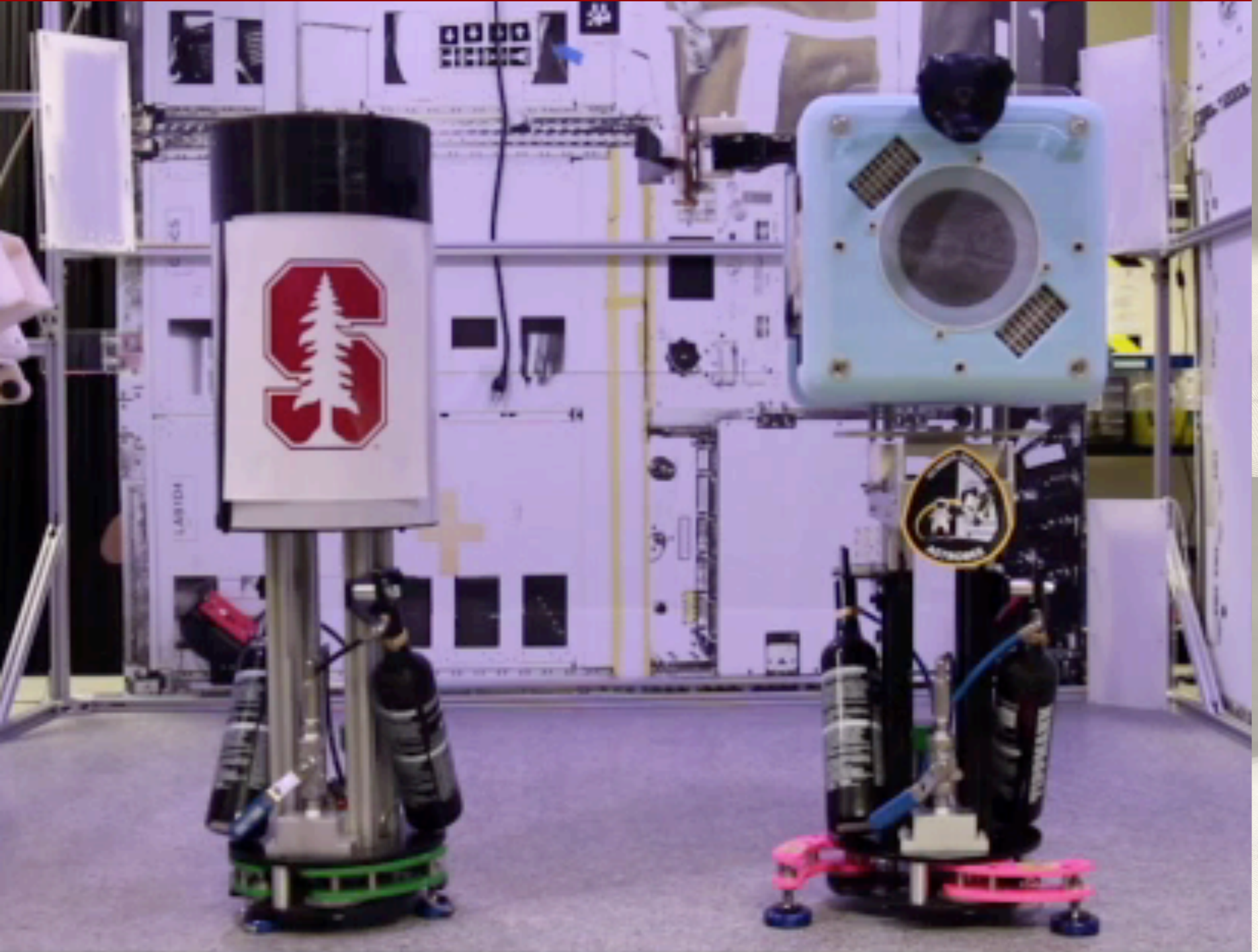
pavone@stanford.edu

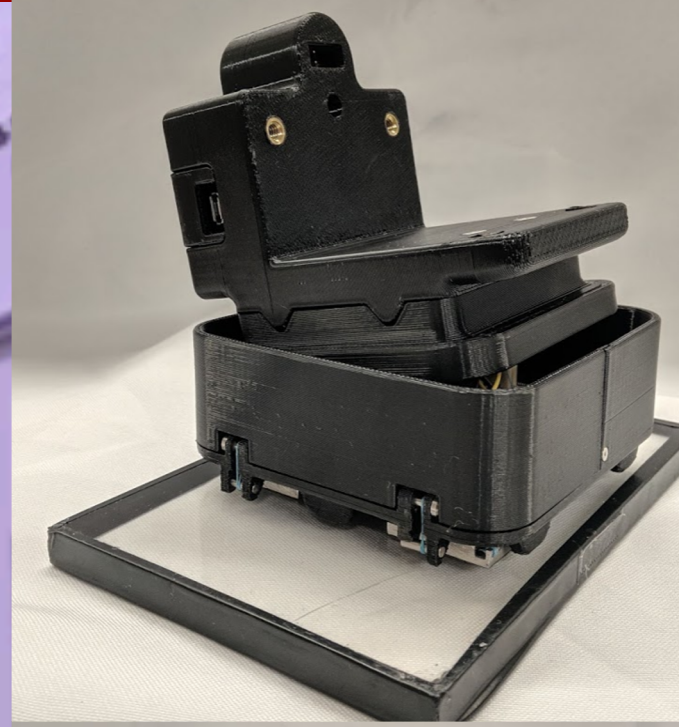


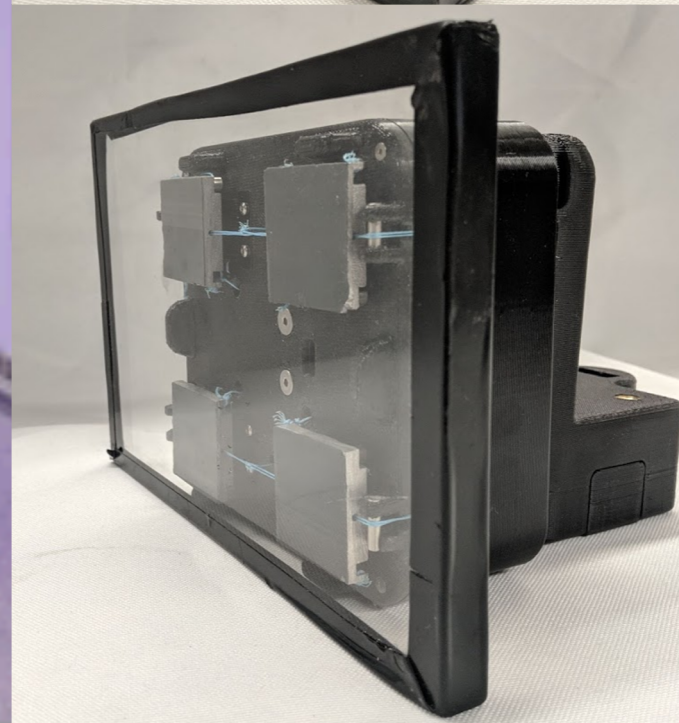


Courtesy of NASA JPL









From reactive to proactive decision making

> “Merging into traffic during rush hour is an exercise in negotiation."
>
> — Google Report, 2016

Proactive decision making:

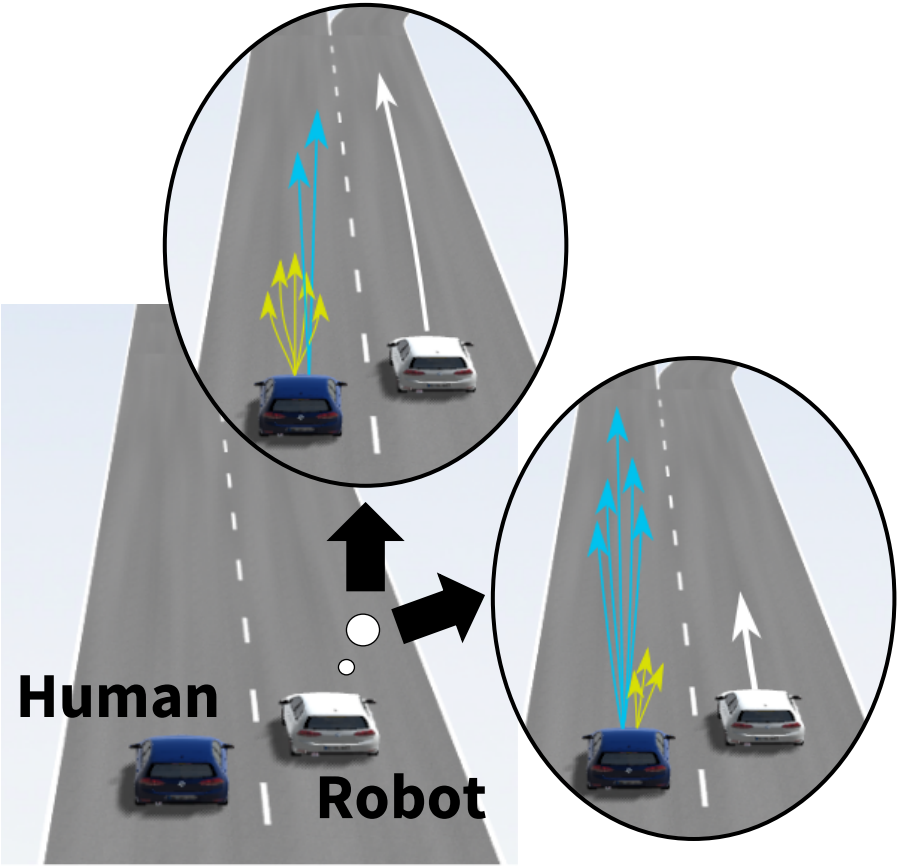
1. proactively interact with other agents to infer their intents, while concurrently
2. exploiting this information to take actions that account for agent responses

High-level considerations

Approach: *model-based* decision making for pairwise interaction

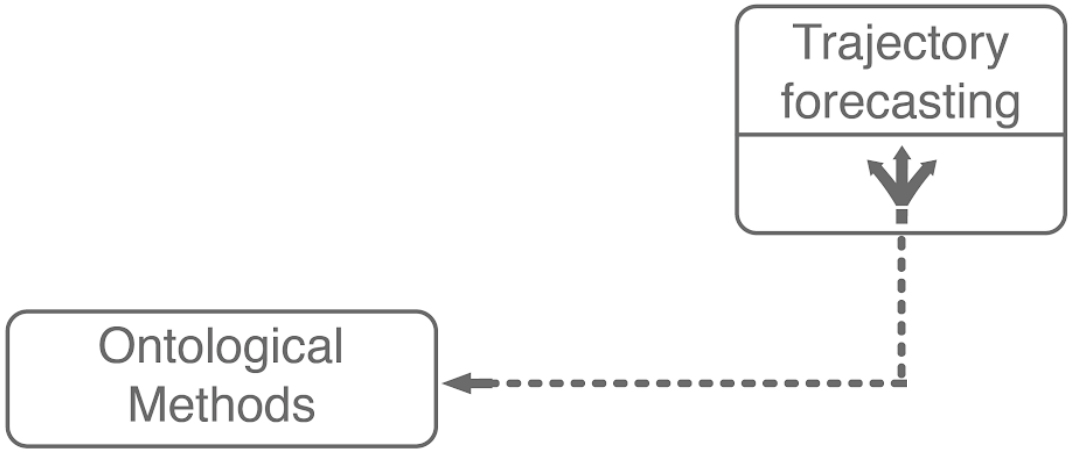
Key considerations:

1. Conditional prediction of human intent
2. Time scales on the order of ~1s
3. Uncertainty is generally multimodal
4. History-dependent predictions
5. Interpretability



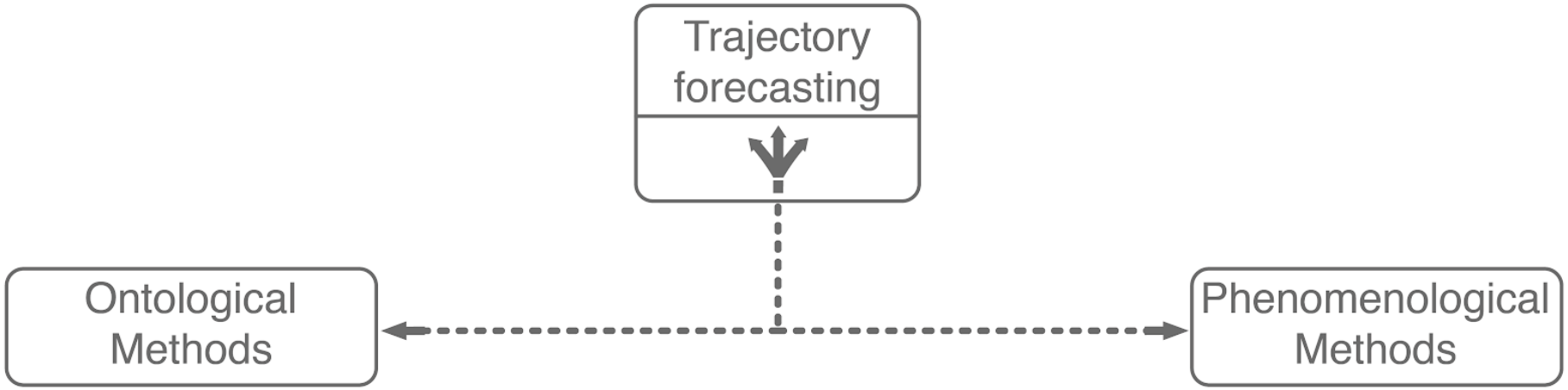
Data-driven probabilistic modeling

[Schmerling, Leung, Vollprecht, Pavone, ICRA ‘18 & Ivanovic, Schmerling, Leung, Pavone, IROS ‘18]



Data-driven probabilistic modeling

[Schmerling, Leung, Vollprecht, Pavone, ICRA '18 & Ivanovic, Schmerling, Leung, Pavone, IROS '18]



Phenomenological approach: can we learn action distributions directly from experience, without reasoning about motivations?

Generative model of human action distributions

[Schmerling, Leung, Vollprecht, Pavone, ICRA '18 & Ivanovic, Schmerling, Leung, Pavone, IROS '18]

Generative model of human action distributions conditioned on:

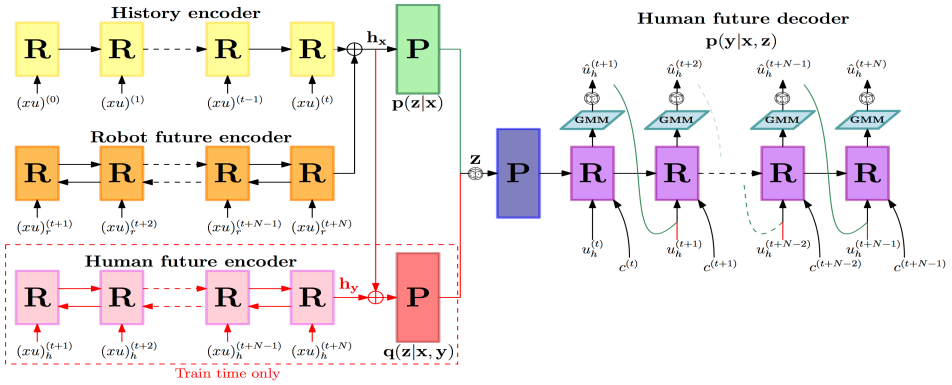
- Joint interaction history
- Candidate robot future action sequence

$$U_H^{(t+1)} \sim p(u_H^{(t+1)} \mid x^{(0:t)}, u^{(0:t)}, u_R^{(t+1)})$$

Interaction History

Robot's Next Action

Learnt via a CVAE-based model

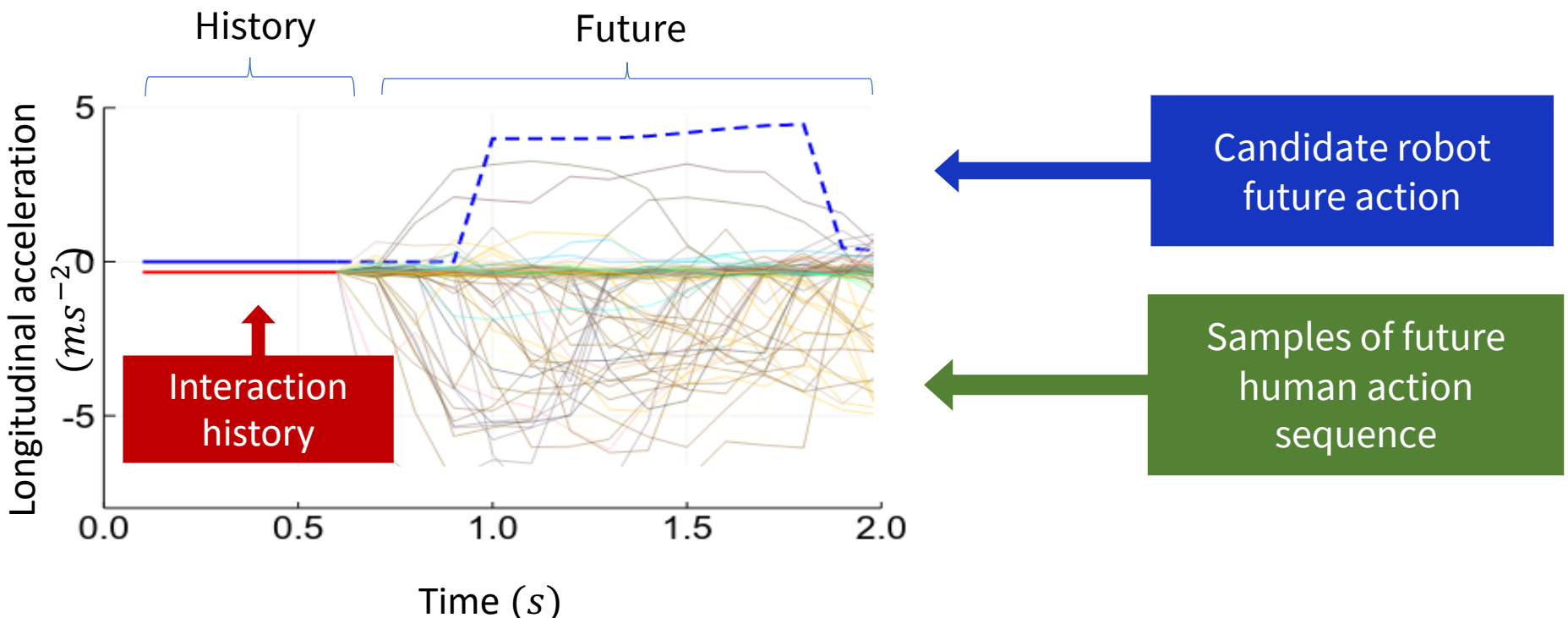


Generative model of human action distributions

[Schmerling, Leung, Vollprecht, Pavone, ICRA '18 & Ivanovic, Schmerling, Leung, Pavone, IROS '18]

Generative model of human action distributions conditioned on:

- Joint interaction history
- Candidate robot future action sequence

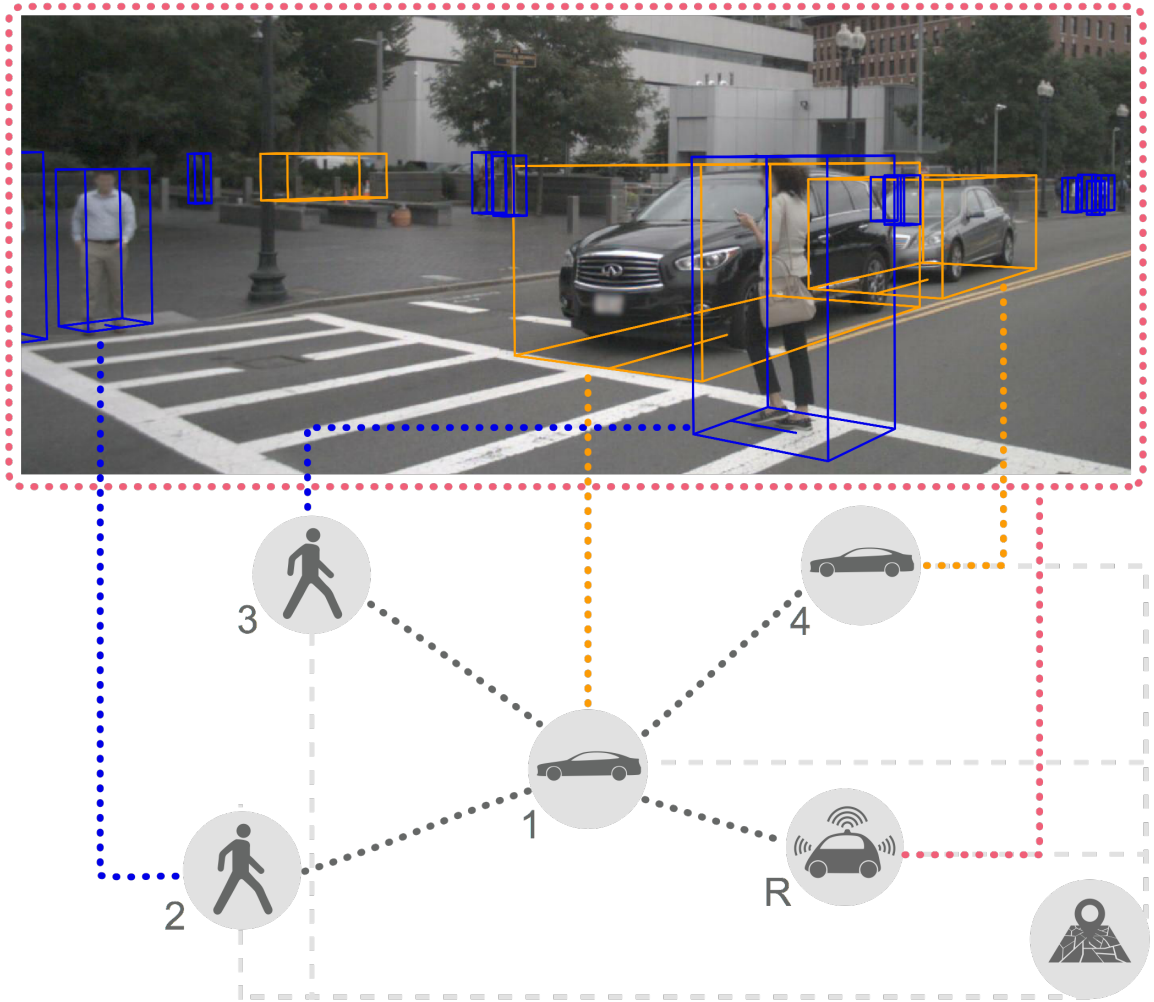


General multi-agent trajectory modeling

[Ivanovic, Pavone, ICCV '19 & Salzmann, Ivanovic, Chakravarty, and Pavone, arXiv '20]

Towards a general multi-agent model that can ingest "everything"

1. Predictions for any kind of agent
2. Accounting of dynamics constraints
3. Conditioning on heterogenous data

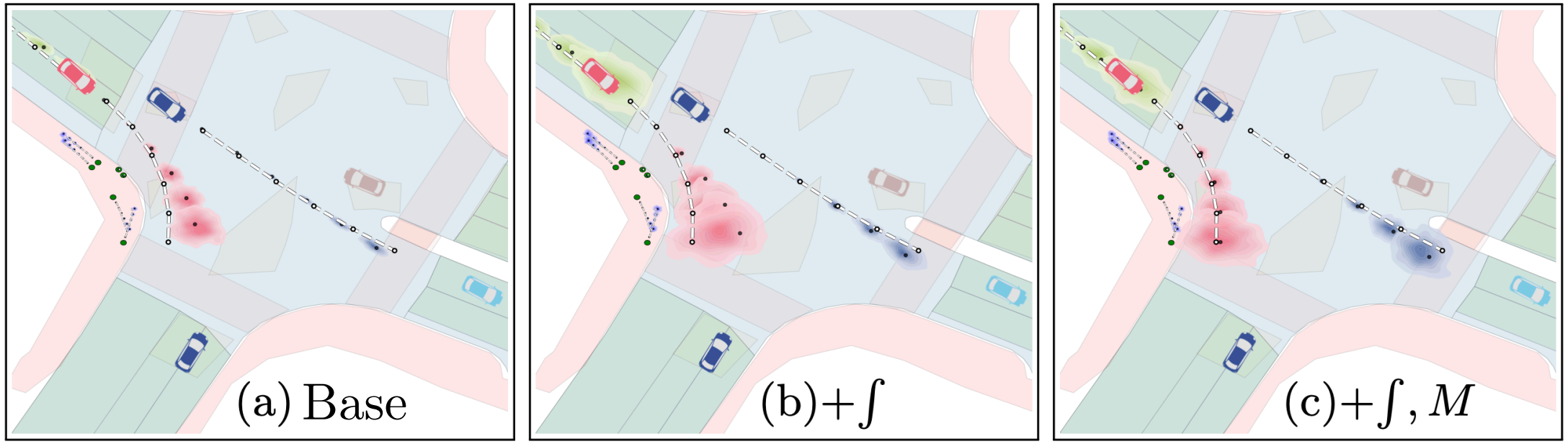


Trajectron++

[Salzmann, Ivanovic, Chakravarty, and Pavone, arXiv '20]

Trajectron++

State-of-the-art generative model that explicitly incorporates agent dynamics and heterogeneous data

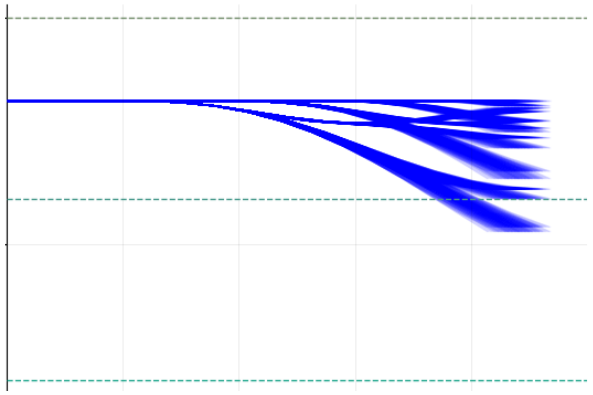


All code, models, and data available at https://github.com/StanfordASL/Trajectron-plus-plus

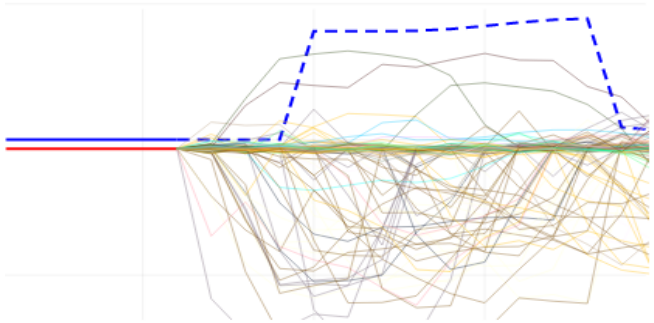
Takeaway message: phenomenological models (and, in particular, deep generative models) achieve state-of-the-art performance

Robot policy construction

- Robot builds an action tree (~4,000 action sequences)



- Model used to sample human responses (>100K!)

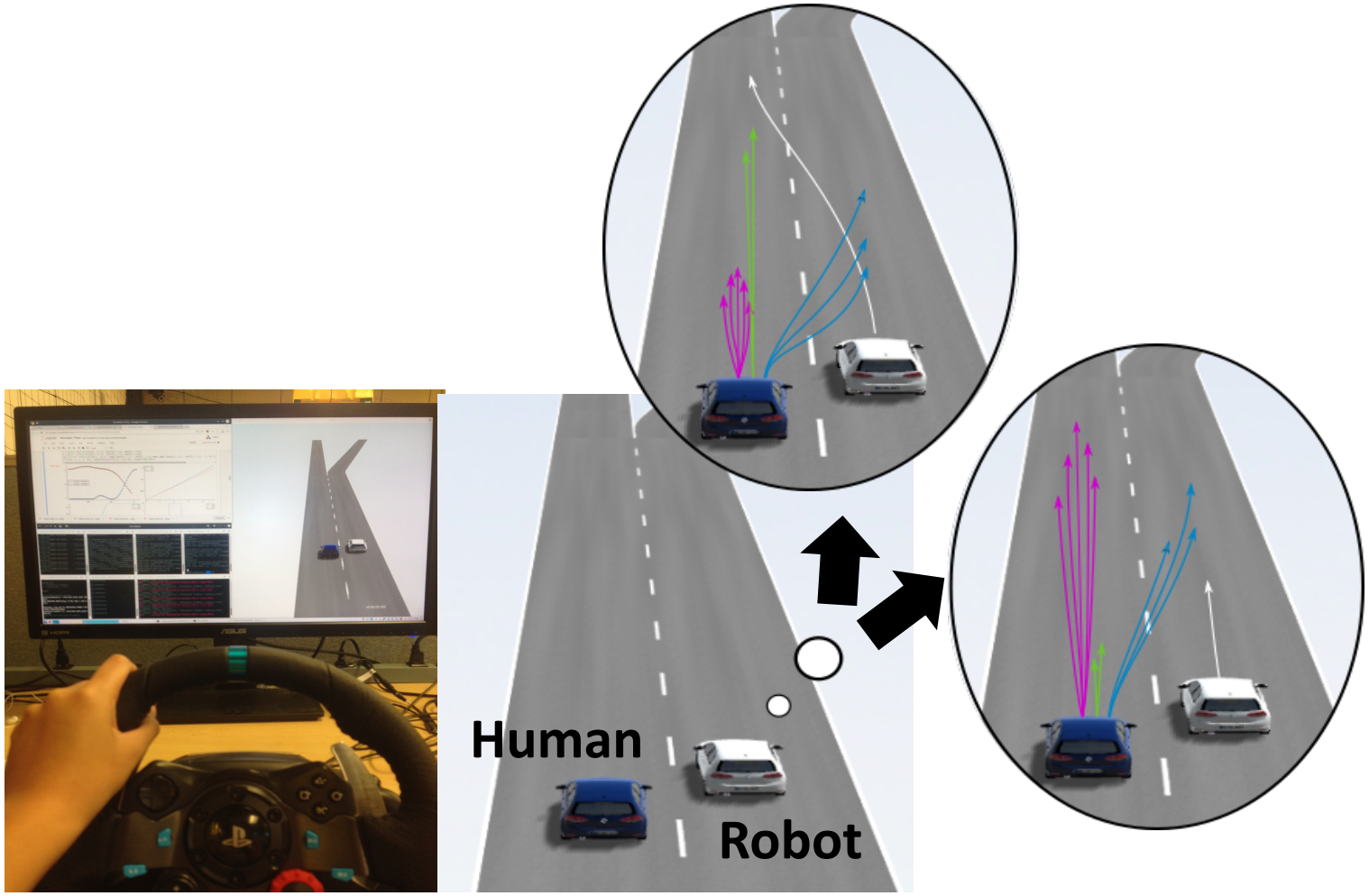


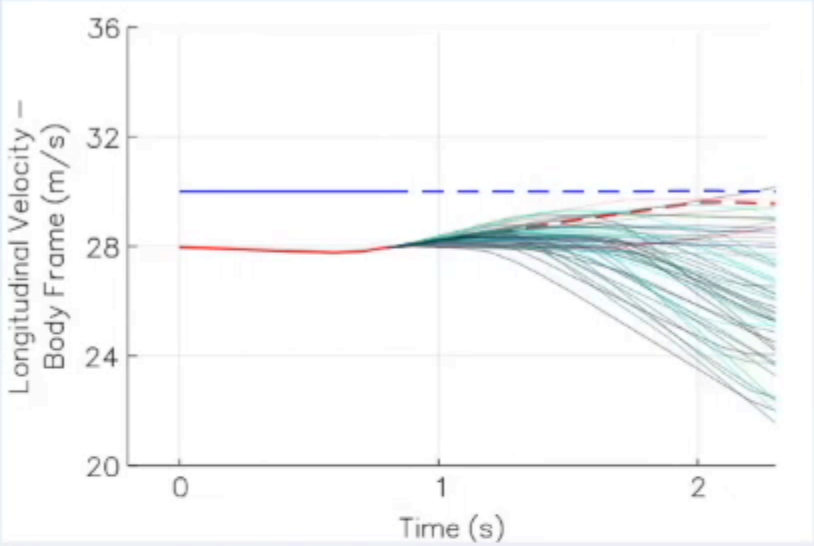
- Action sequences scored via an *aggregate* cost function

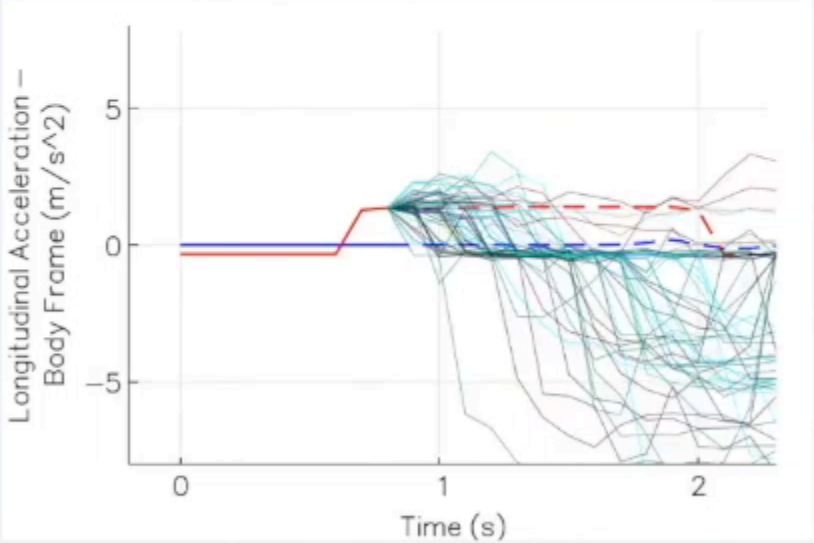
$$\mathbb{E}\left(\sum_i J_i\right) \xrightarrow{?} \rho\left(\sum_i J_i\right)$$

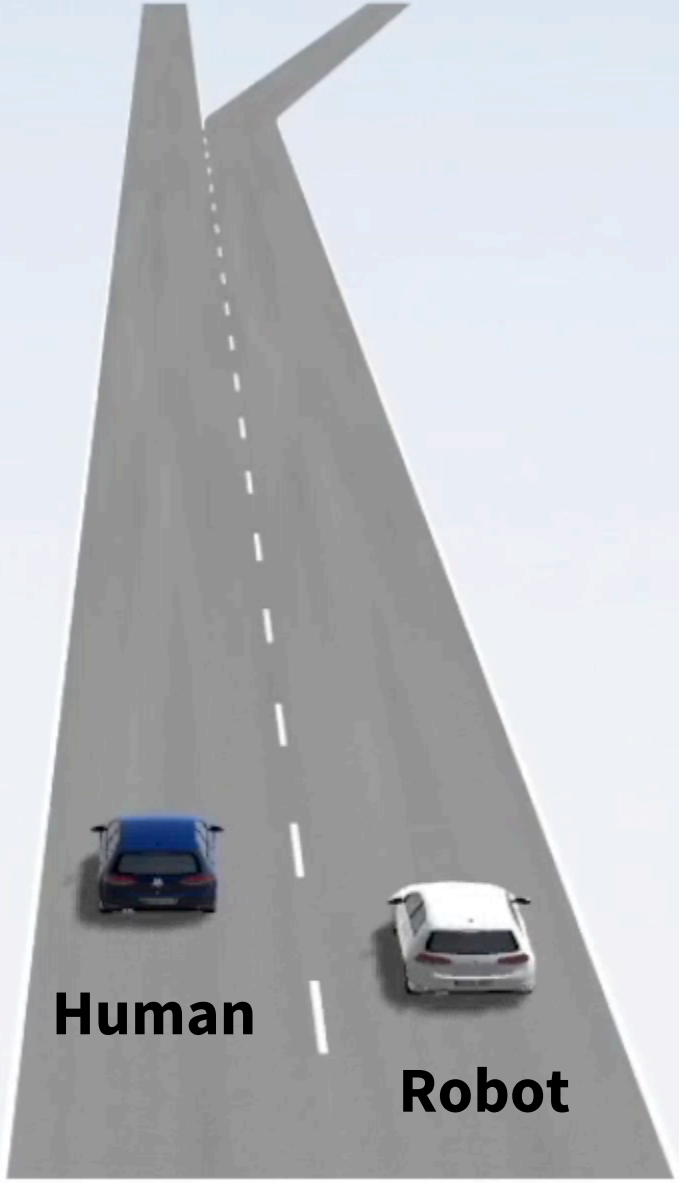
[Majumdar, Pavone, ISRR '17]

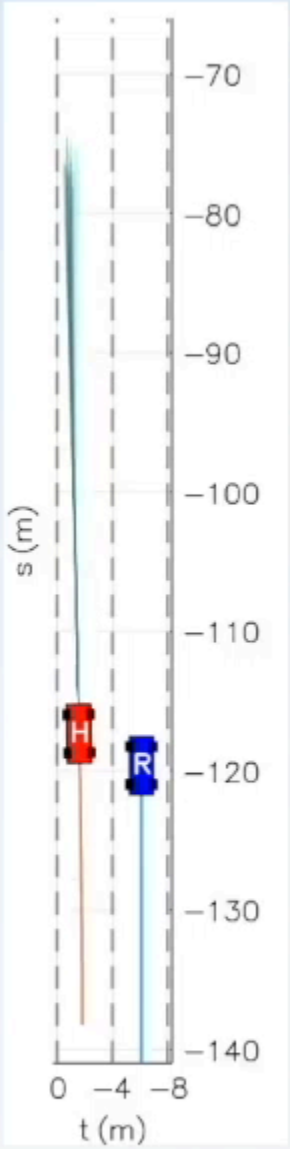
Human-in-the-loop testing



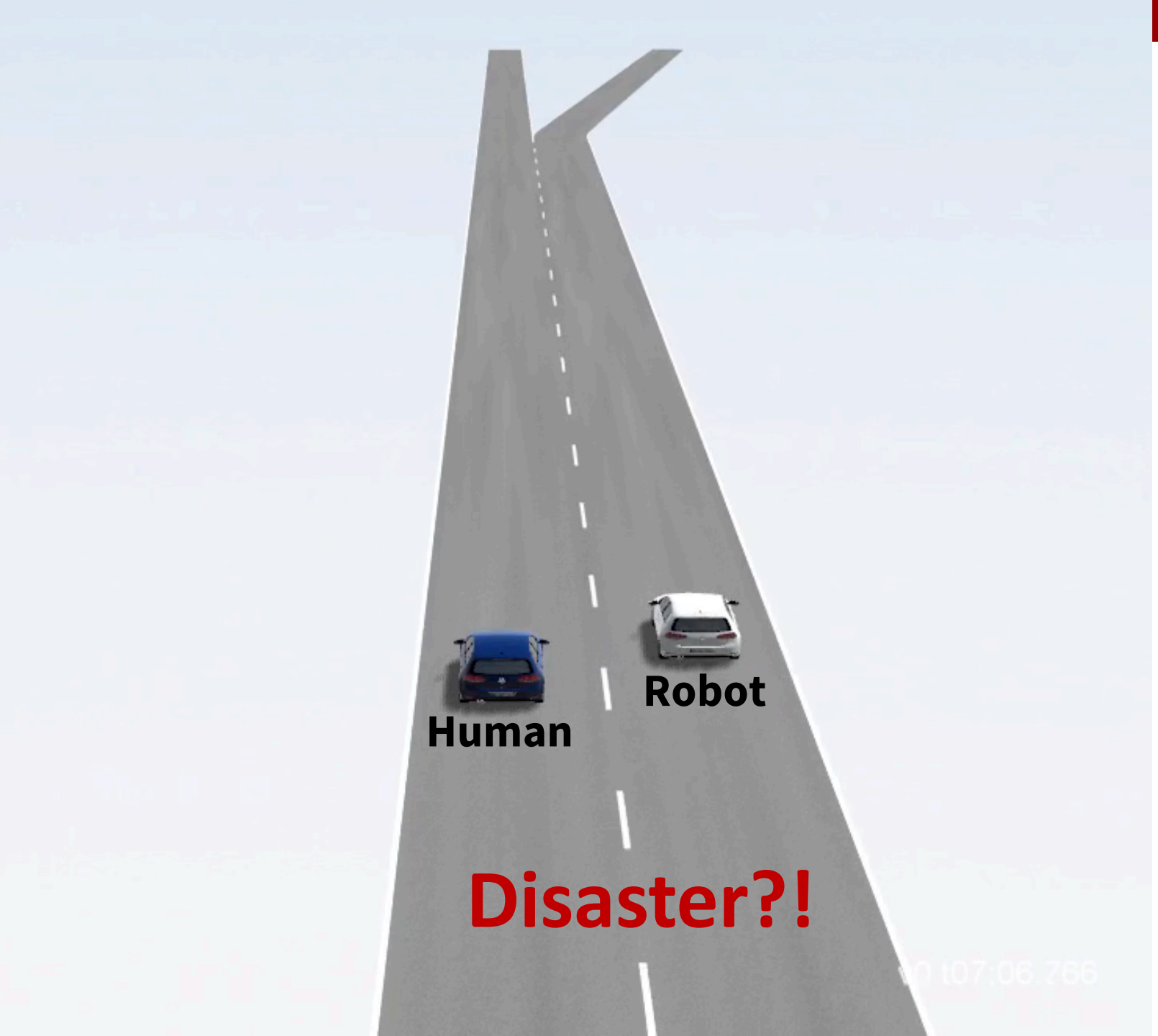

Longitudinal Velocity – Body Frame (m/s)
36
32
28
24
20
0
1
2
Time (s)


Longitudinal Acceleration – Body Frame (m/s^2)
5
0
−5
0
1
2
Time (s)


Human
Robot

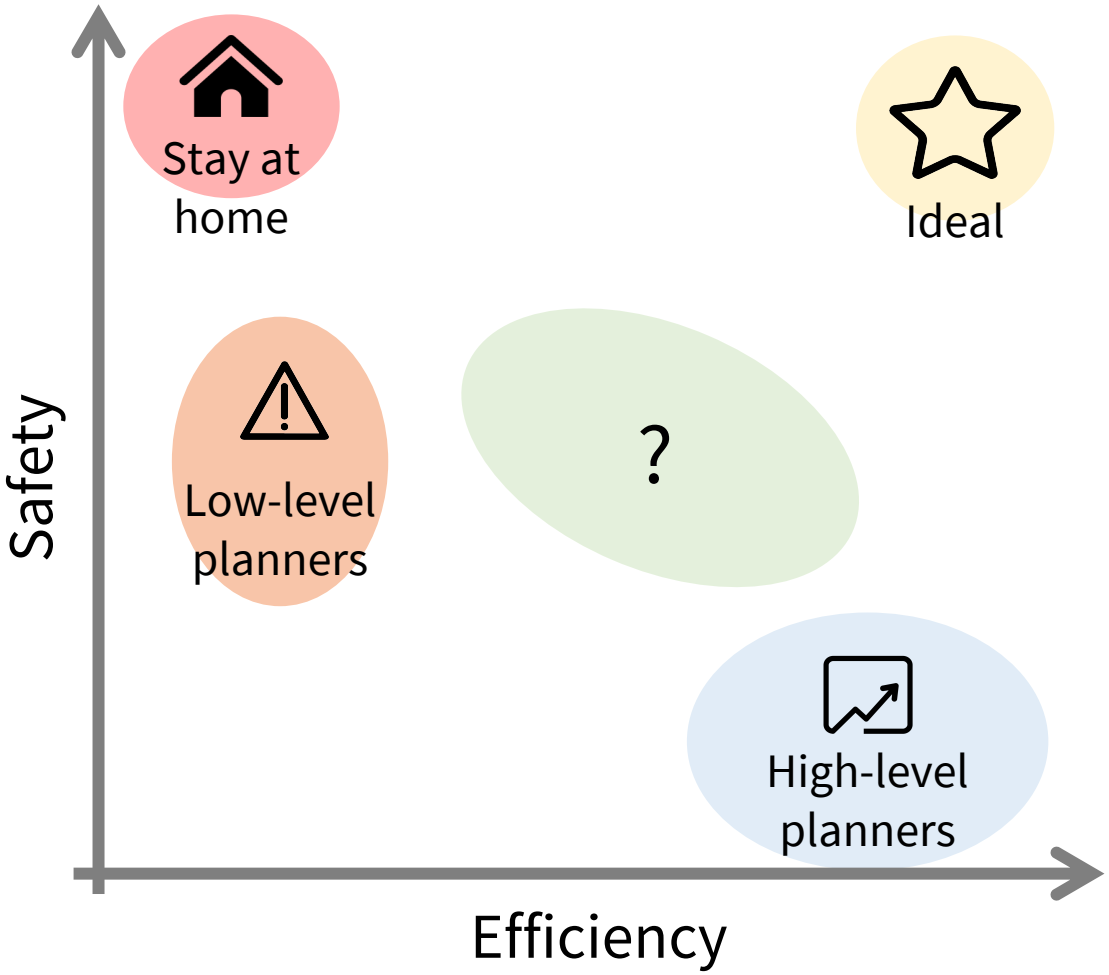

s (m)
−70
−80
−90
−100
−110
−120
−130
−140
H
R
0
−4
−8
t (m)


0
−4
−8



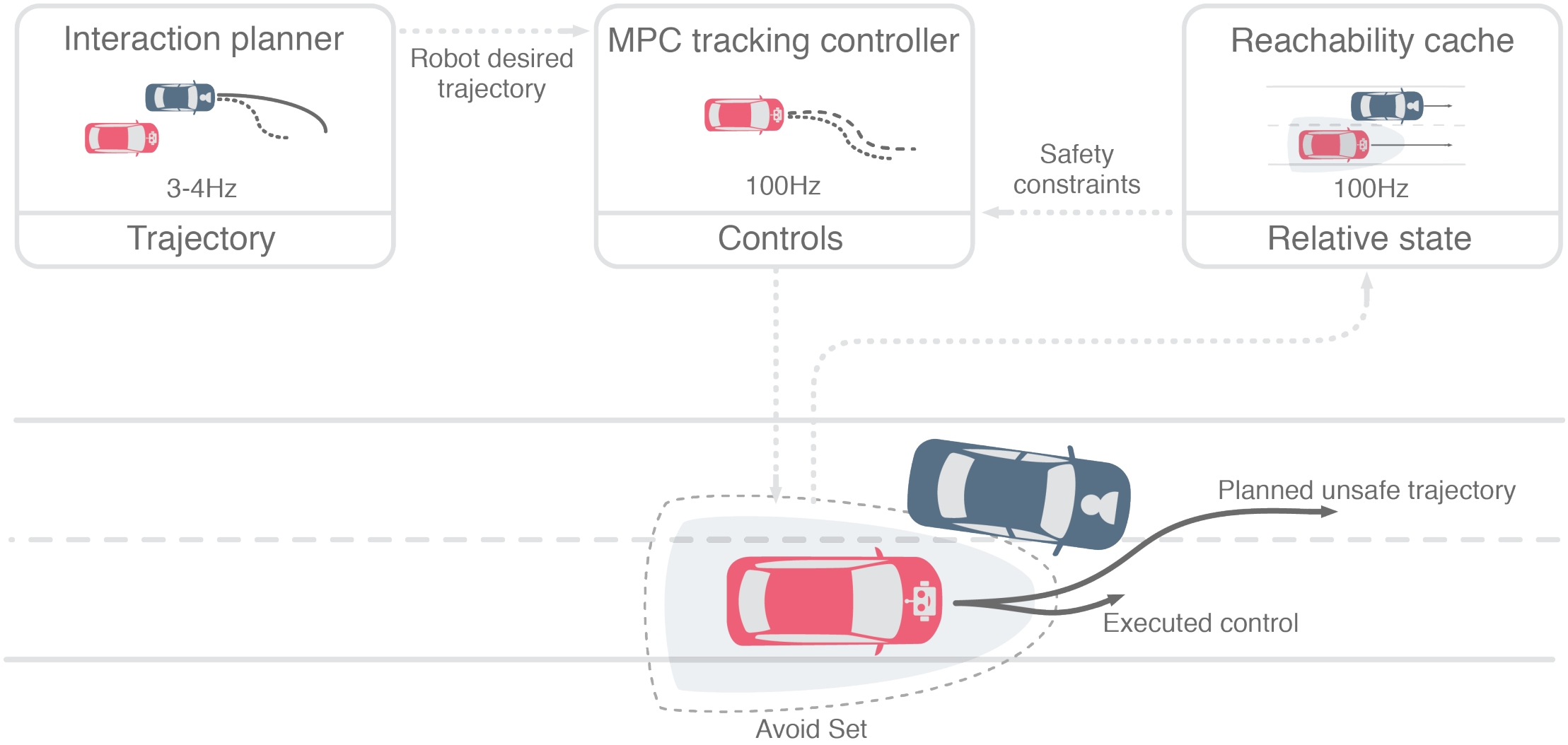
- Probabilistic model may “get it wrong”
- Incorporating collision avoidance as a penalty can cause conflicting objectives
- Replanning at 3Hz is ultimately too slow to ensure safety

How to integrate safety assurance within a probabilistic, performance-centric planning framework?

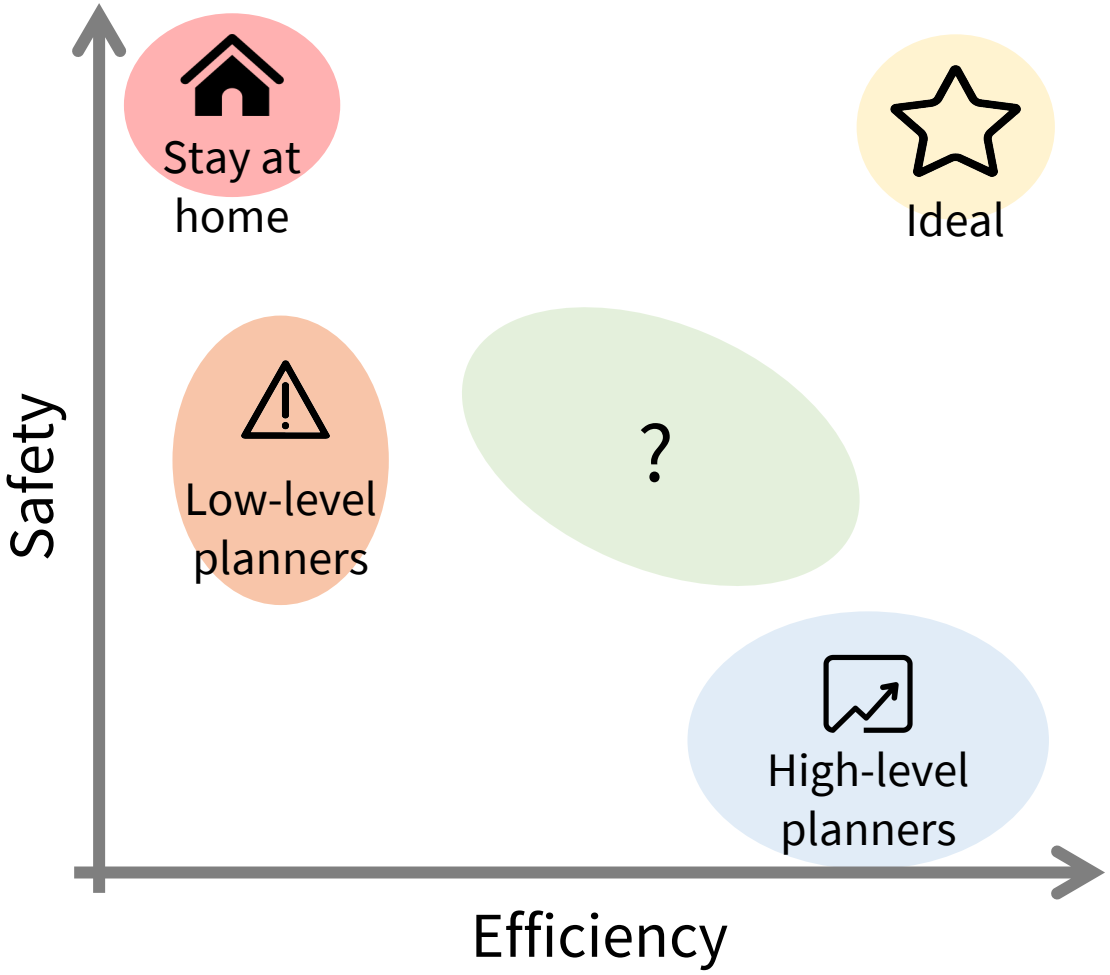


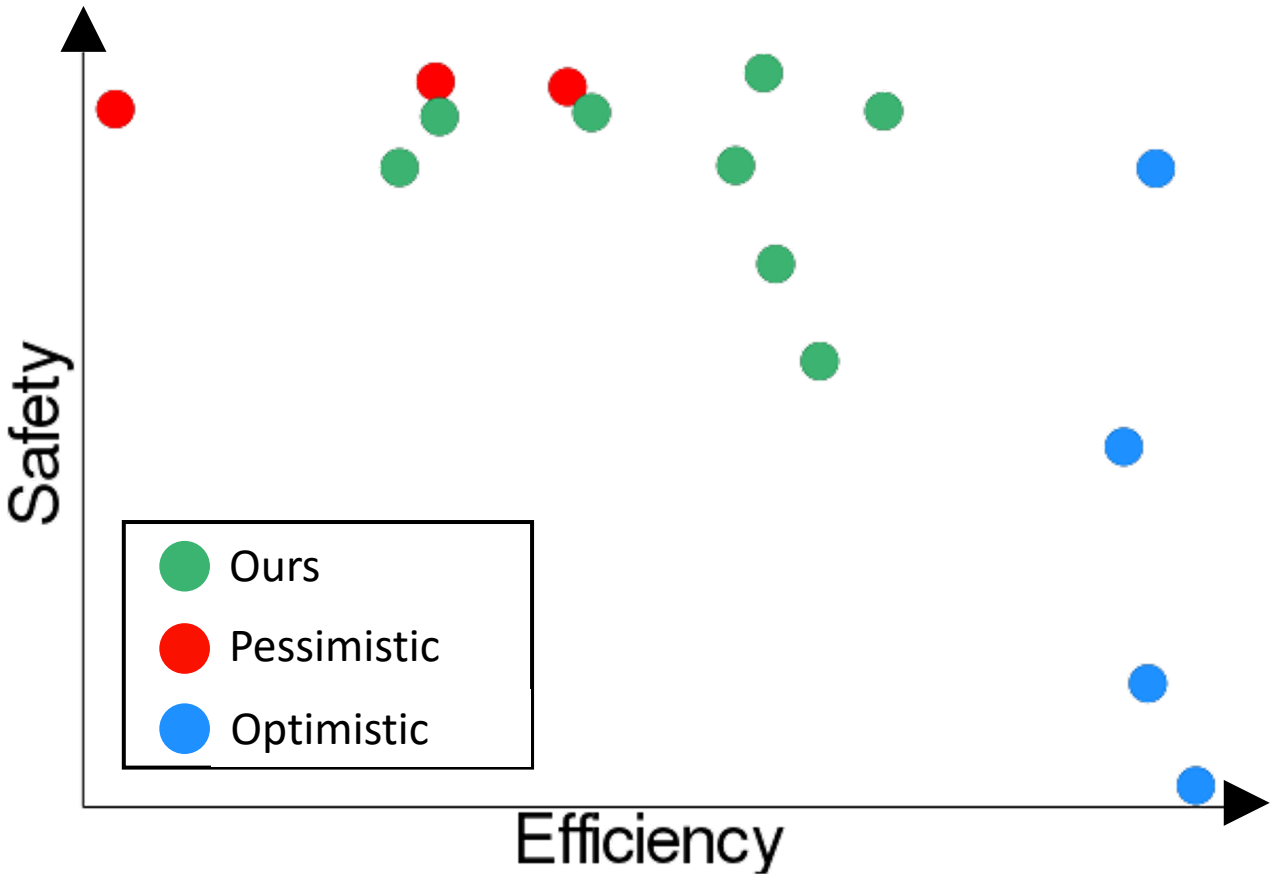
Full decision-making and control stack

[Leung, Schmerling, Chen, Talbot, Gerdes, Pavone, ISER '18 and IJRR '20]



On the safety vs. efficiency trade-off





Takeaway message: safe and efficient
interaction-aware decision making requires
principled blending of probabilistic and adversarial planning

Efficient representations for decision making

[Ivanovic, Elhafsi, Rosman, Gaidon, Pavone, in preparation]

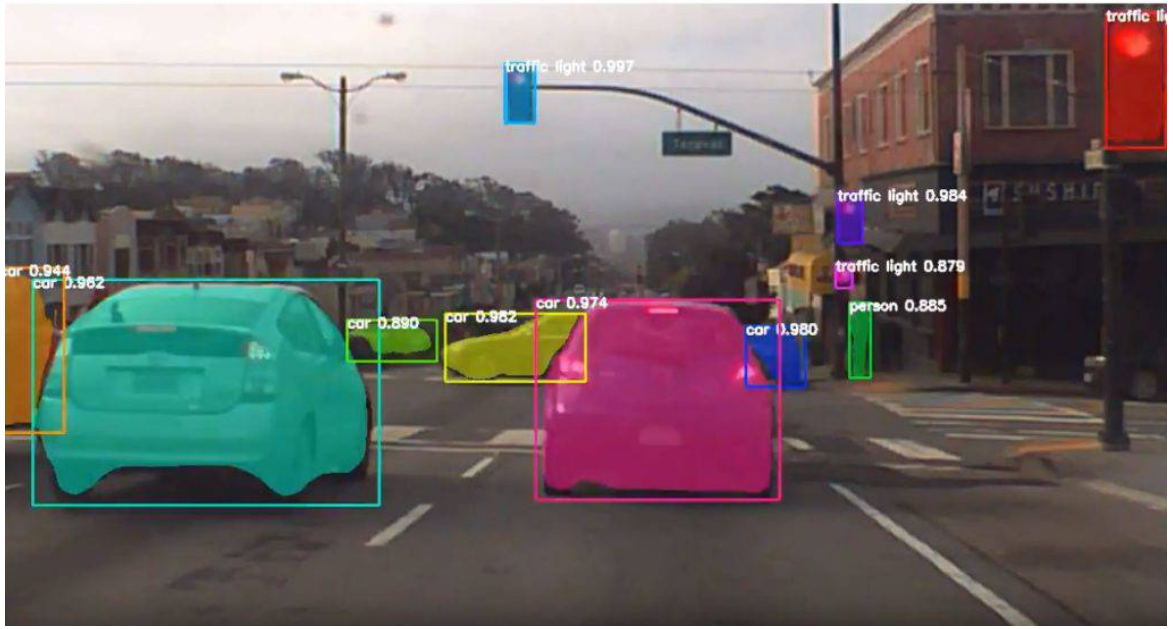
- Prediction of individual trajectories good for evaluation…
- …but difficult to use in downstream decision making
- Idea: reason about prediction representations cognizant of downstream control applications

$$x^{(t+1)} = A^{(t)}x^{(t)} + B^{(t)}u_{\mathrm{R}}^{(t)} + C^{(t)}w \qquad (w \sim \mathcal{N}(0,1))$$

Integrating perception & trajectory forecasting

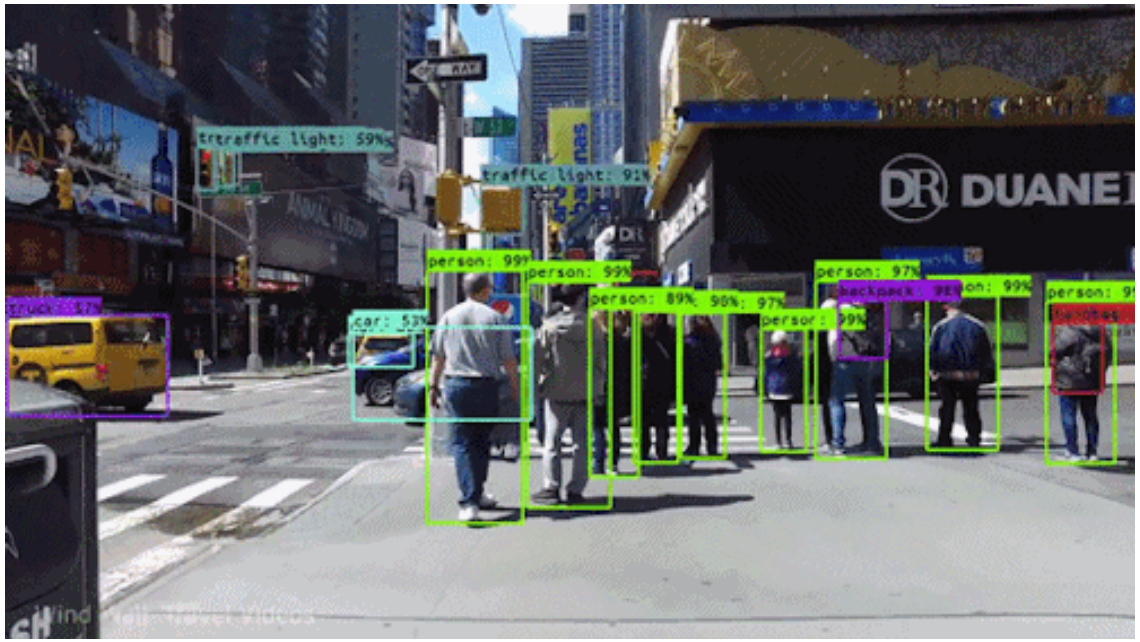
Uncertainty Propagation

Uncertainty in perception should lead to uncertainty in prediction



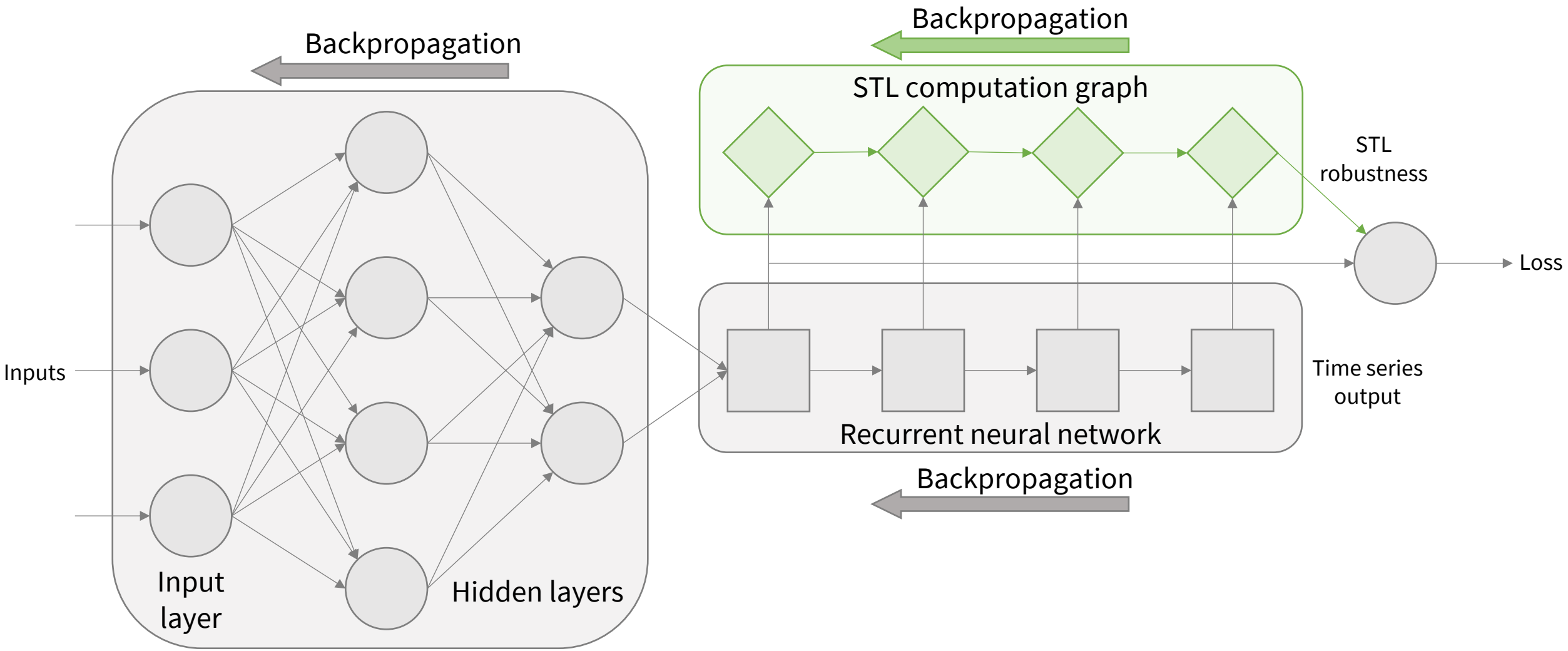
Robustness to Perception

Object classification is noisy, prediction models should be robust to this



Adding structure into predictions

[Leung, Arechiga, Pavone, IVS ’19 – patent pending]



Takeaway message: next-generation trajectory forecasting methods should account for downstream control applications and structure

Conclusions

- Generative models becoming SoA tool for trajectory prediction

Multimodal Deep Generative Models for Trajectory Prediction: A Conditional Variational Autoencoder Approach

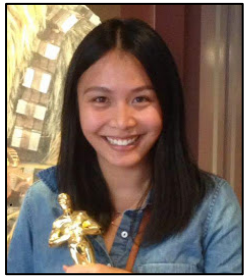
Boris Ivanovic*, Karen Leung*, Marco Pavone

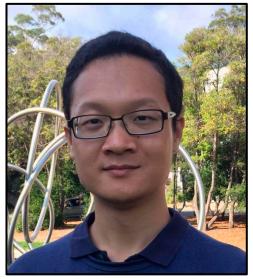
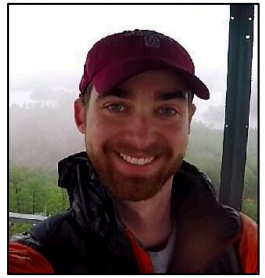
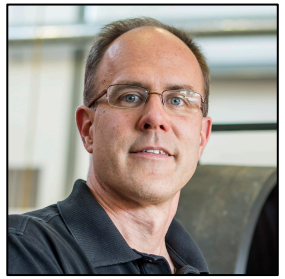
Coming soon!

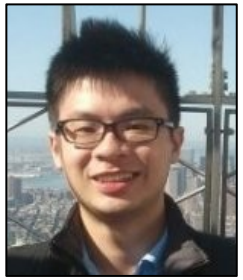
- …but quite difficult to integrate within autonomy stack
  1. reason about prediction representations cognizant of downstream control applications
  2. tighter integration between perception and trajectory prediction
  3. Add logical structure to models

All code available at: https://github.com/StanfordASL

Acknowledgements




pavone@stanford.edu

http://asl.stanford.edu/

https://twitter.com/StanfordASL