Яндекс

ML-планировщик

Как мы делаем машины автономными

Никита Семенов

Intro

12 лет в машинном обучении

Отвечаю за ML

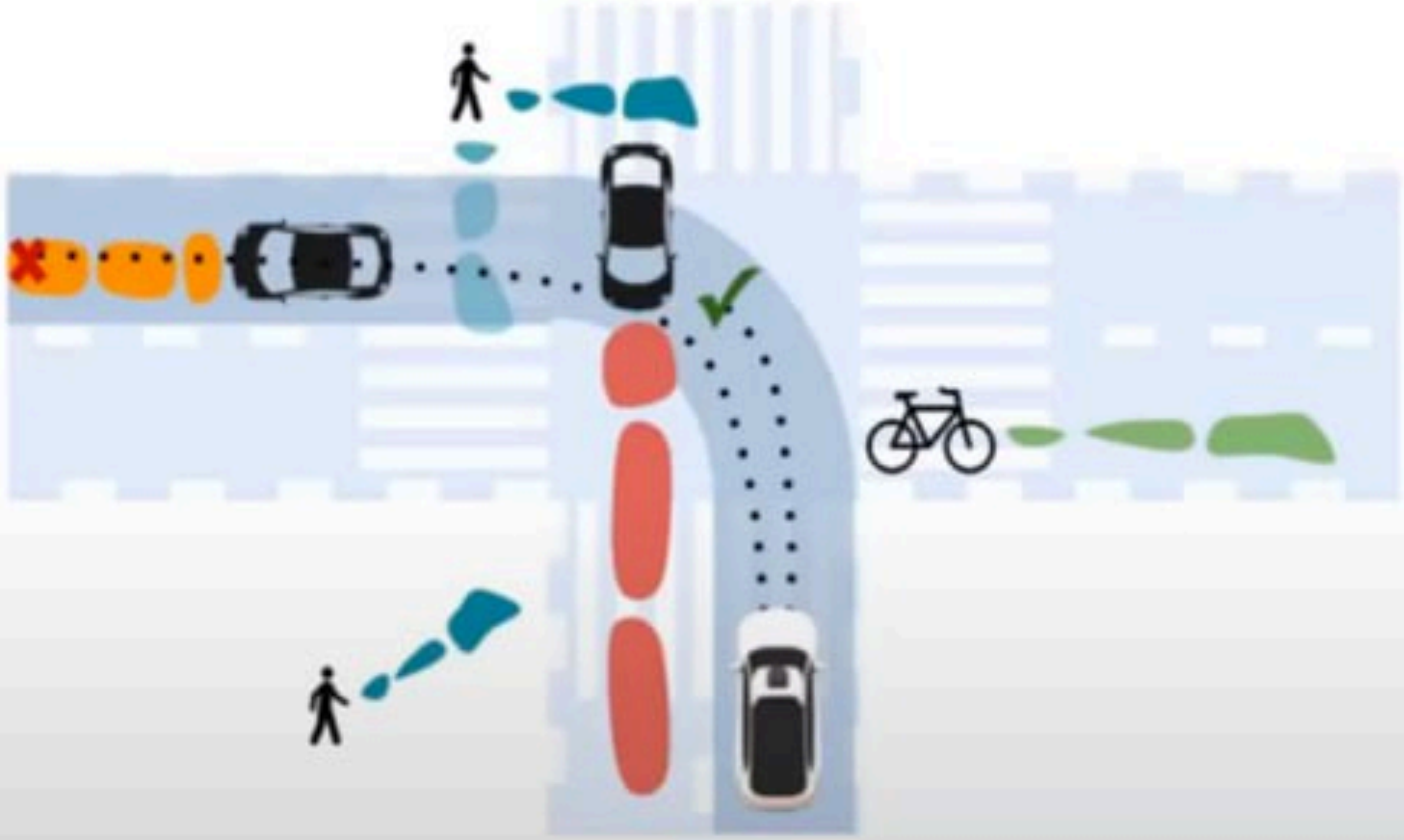
Беспилотные машины

Компоненты пайплайна

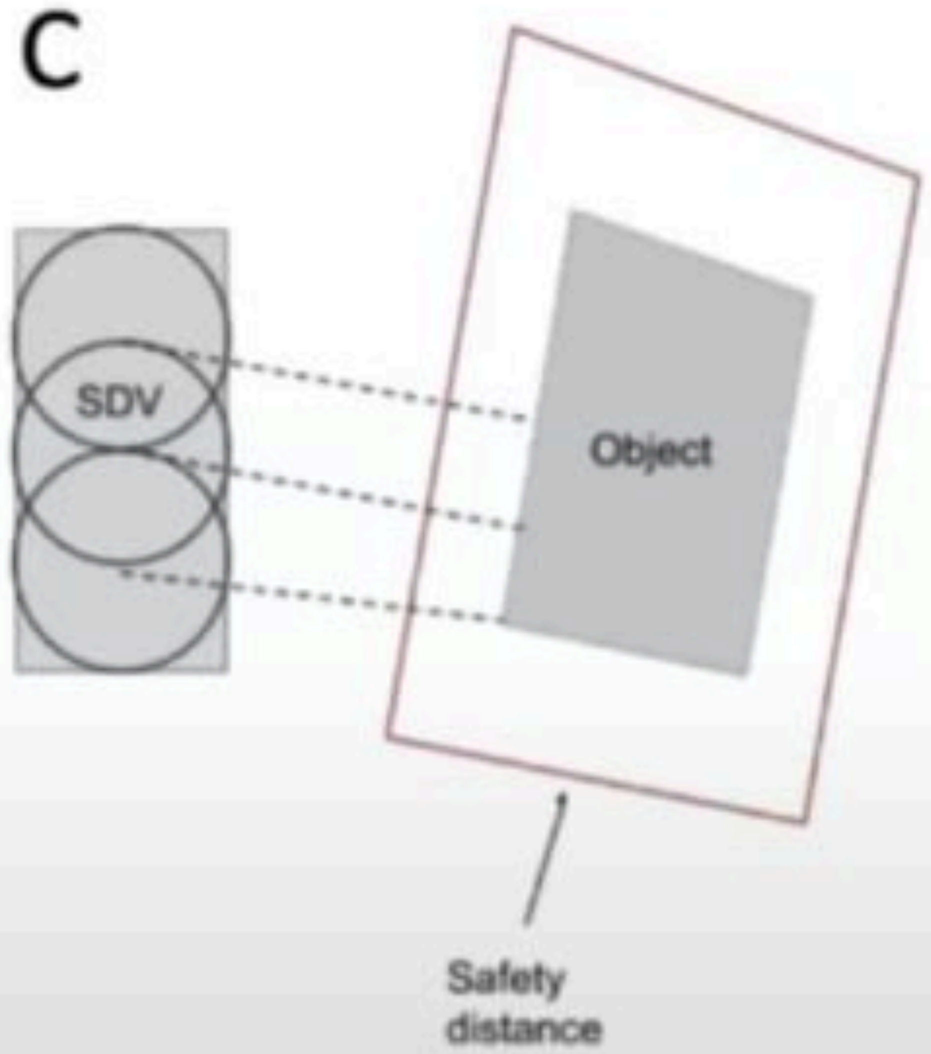
Intermediate representations

Bounding Boxes

Occupancies



Affordances



- Simple, interpretable
- Coarse shapes
- Requires post-processing

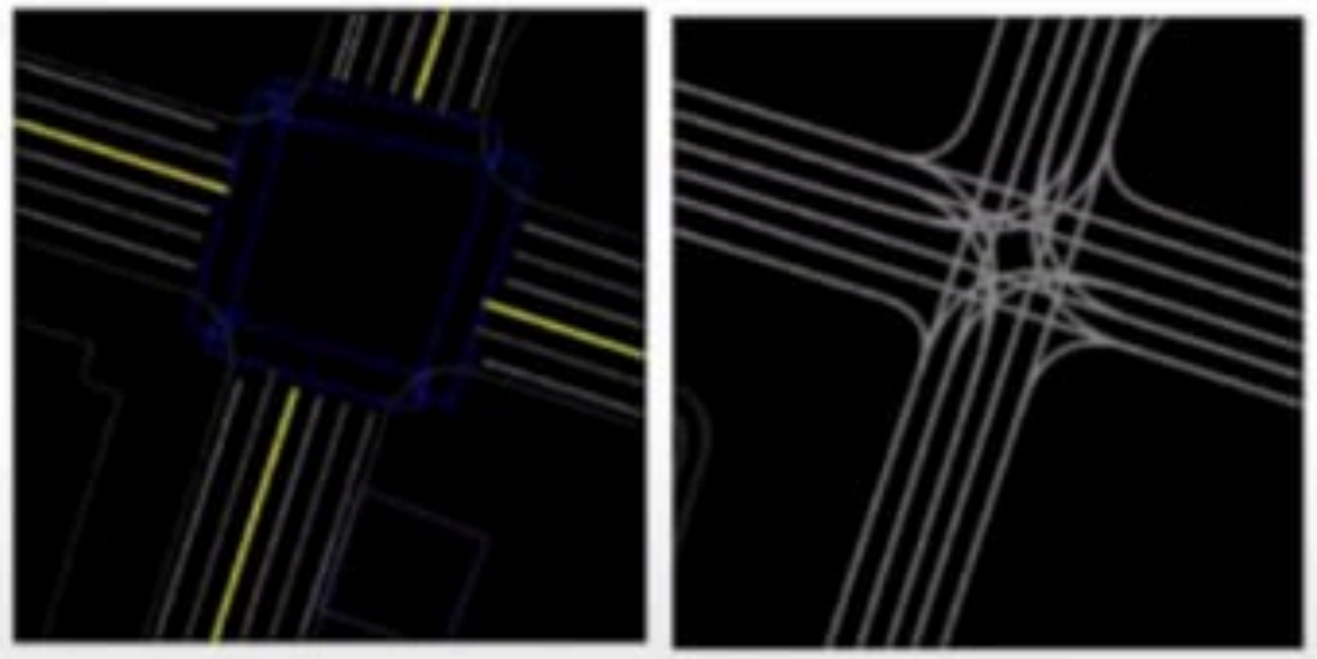
- No post-processing
- Supports complex shapes

- Lightweight representation
- Requires feature engineering

HD Maps

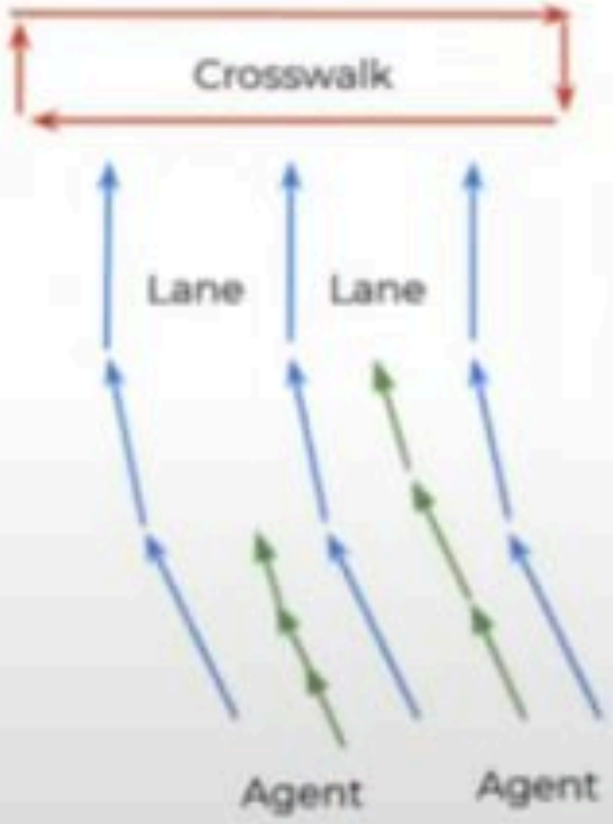
- Useful prior on where the SDV can drive, how to reach its destination, traffic rules, etc.

Rasterization



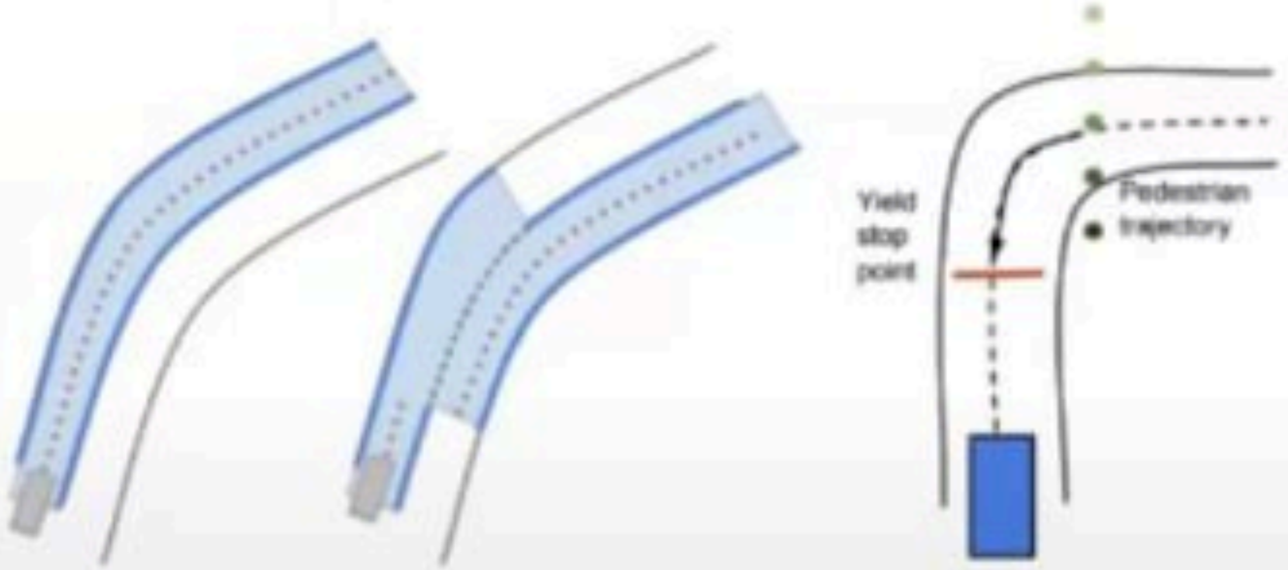
- Simple, easy to process
- Limited receptive fields
- Difficult to incorporate priors

Lane Graph



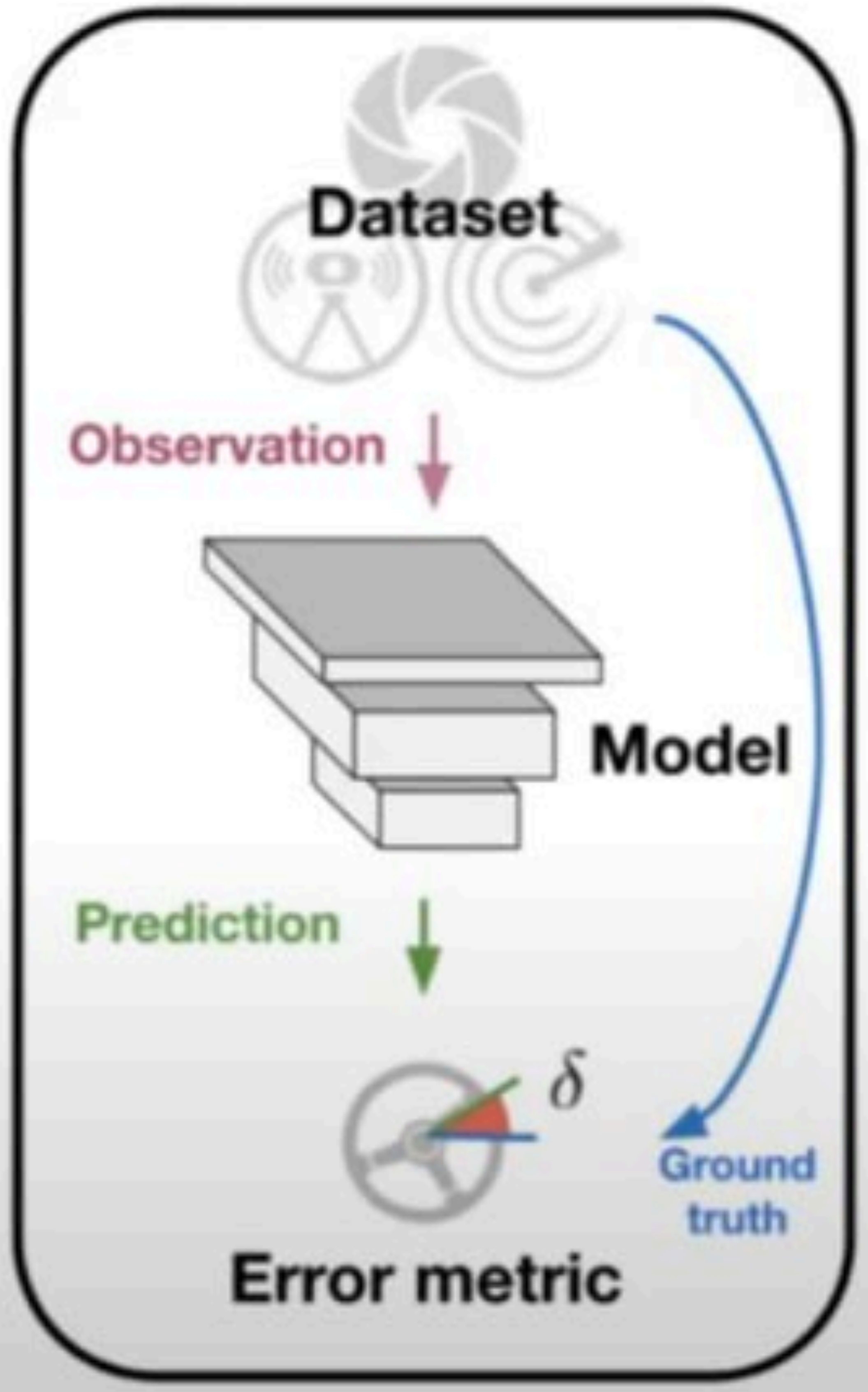
- Larger receptive field
- Complex model architecture

Affordances



- Easy to incorporate priors
- Requires feature engineering

Behavior cloning



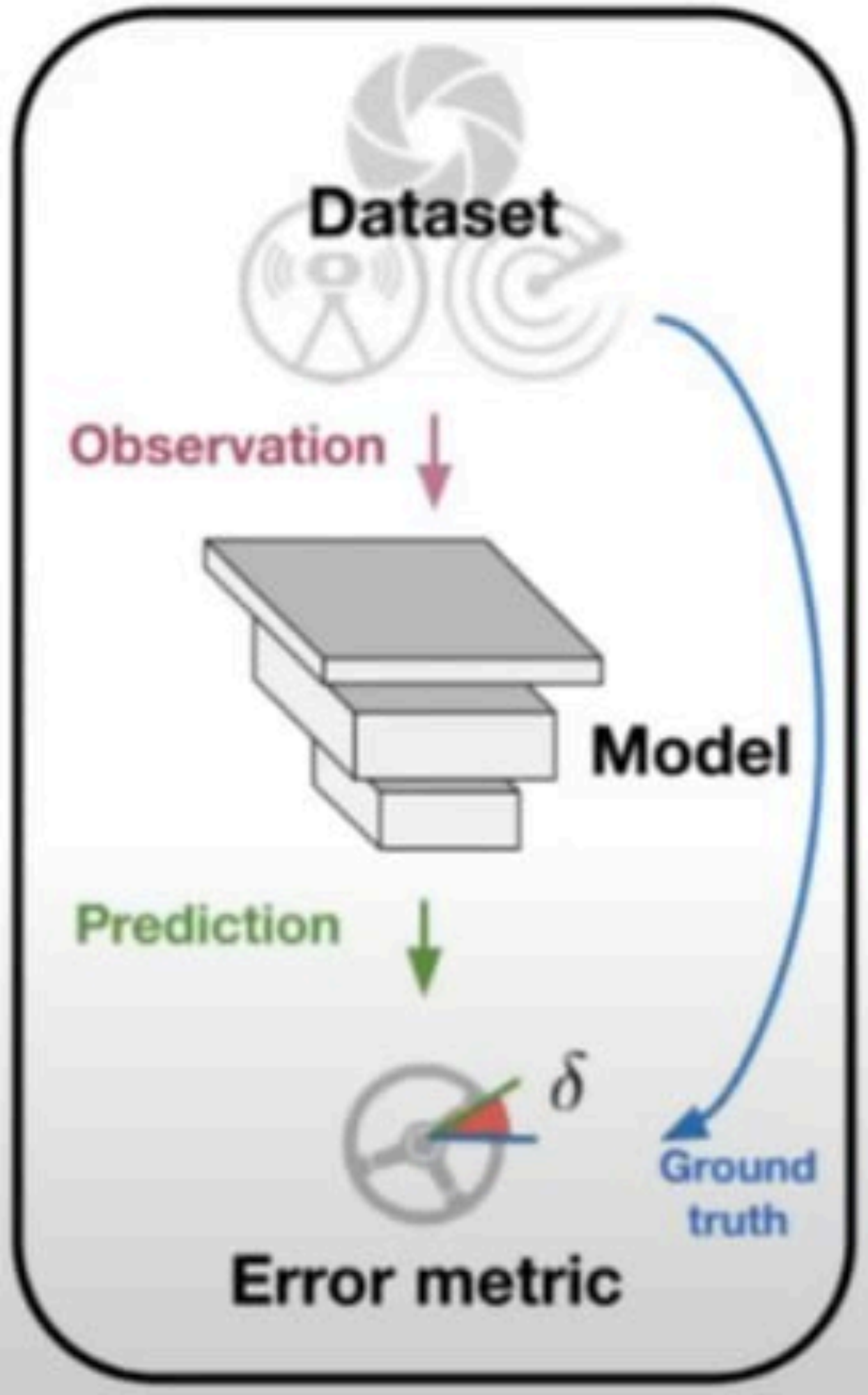
Open-loop learning:

- Policy's actions are never executed

Behavior cloning:

- Policy observes expert states
- Supervised to imitate expert actions:
  - L2 loss between policy and expert actions[1, 2]
  - Max-margin loss to cost policy trajectories higher than expert trajectories[3, 4]

Behavior cloning



Strengths:

- Simple to implement
- High sample efficiency

Weaknesses:

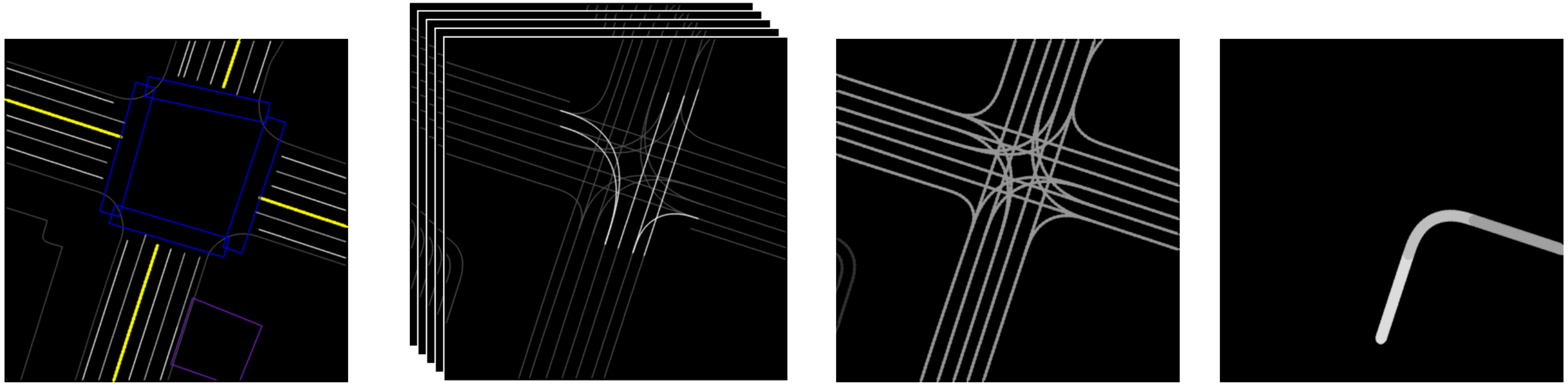
- Distribution shift (training v.s. inference)
- But data augmentation can help[1, 2]

ChauffeurNet

Bansal, Mayank, Alex Krizhevsky, and Abhijit Ogale. "Chauffeurnet: Learning to drive by imitating the best and synthesizing the worst." *arXiv preprint arXiv:1812.03079* (2018).

http://arxiv.org/abs/1812.03079v1

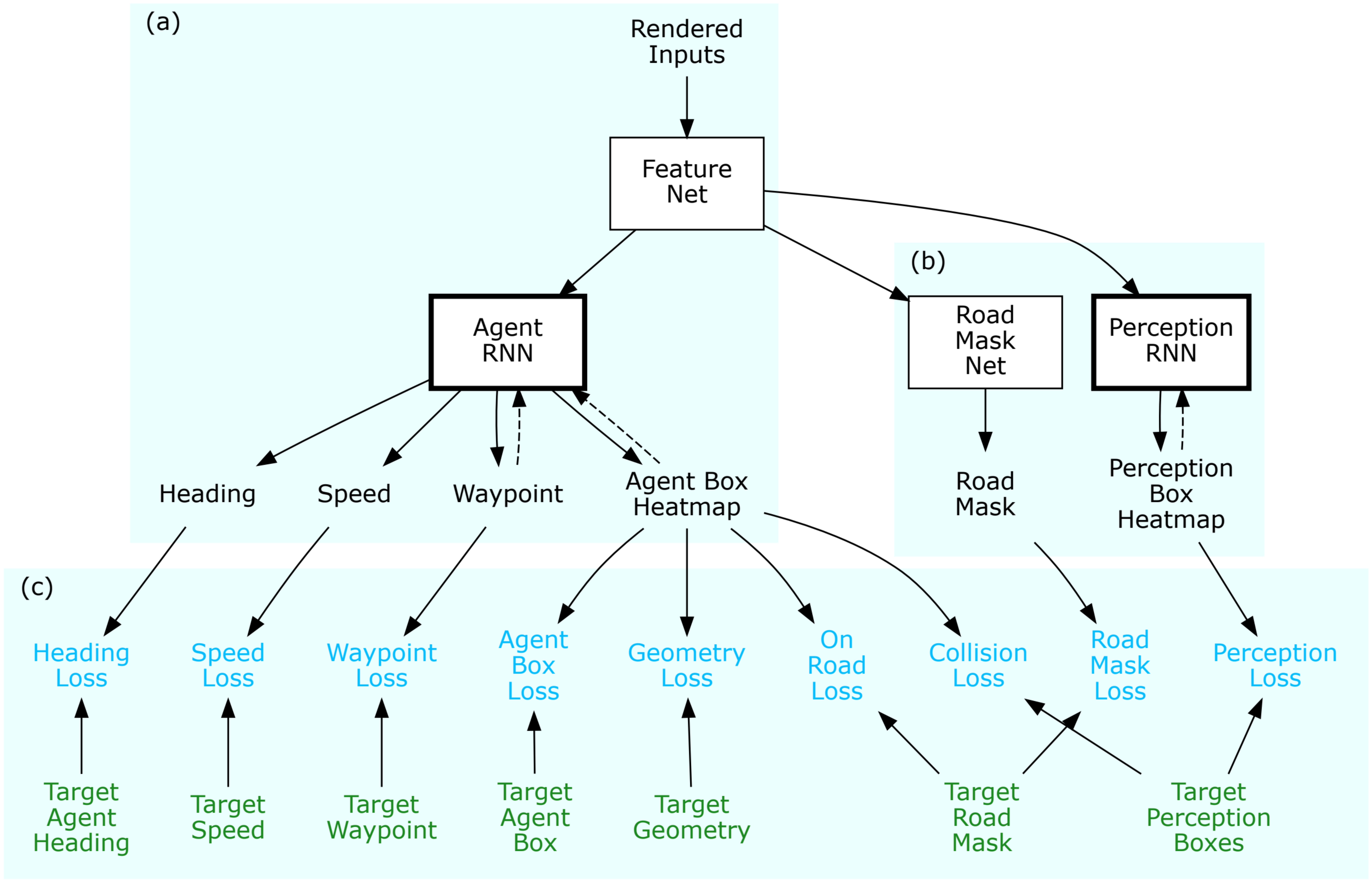
Как видит мир планер?



(a) Roadmap (b) Traffic Lights (c) Speed Limit (d) Route

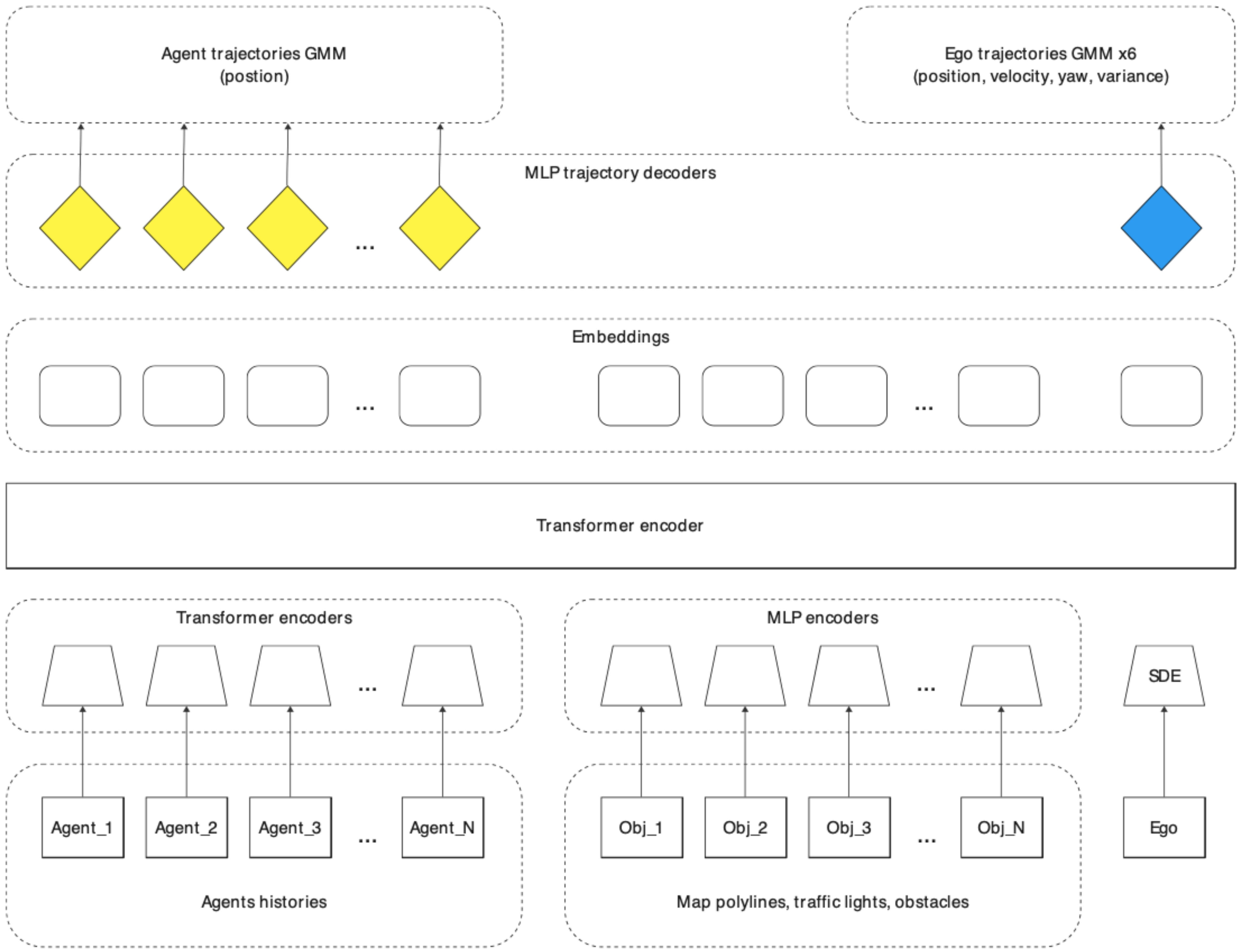
(e) Current Agent Box (f) Dynamic Boxes (g) Past Agent Poses (h) Future Agent Poses

Chauffeur Net

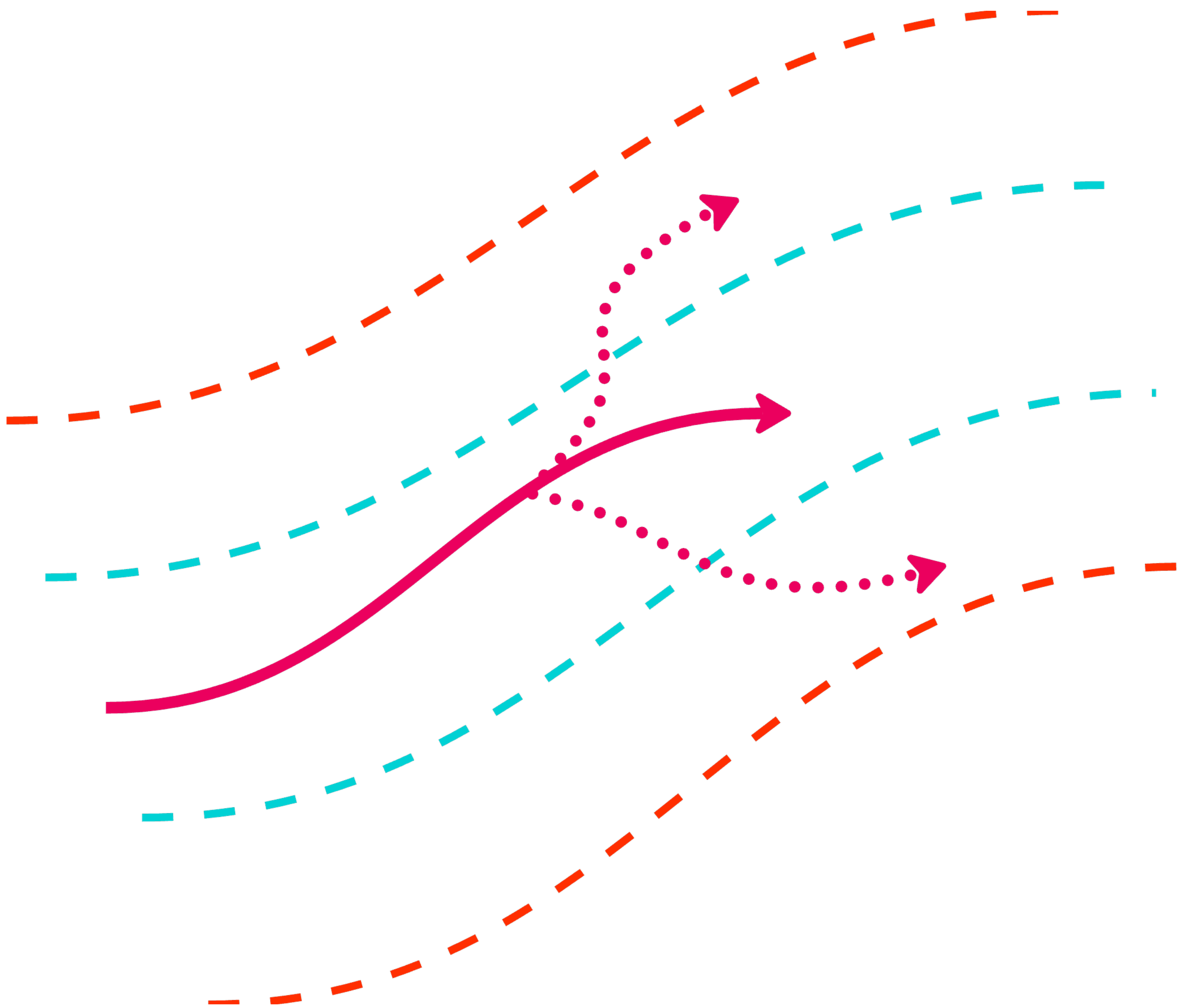


PlanTF

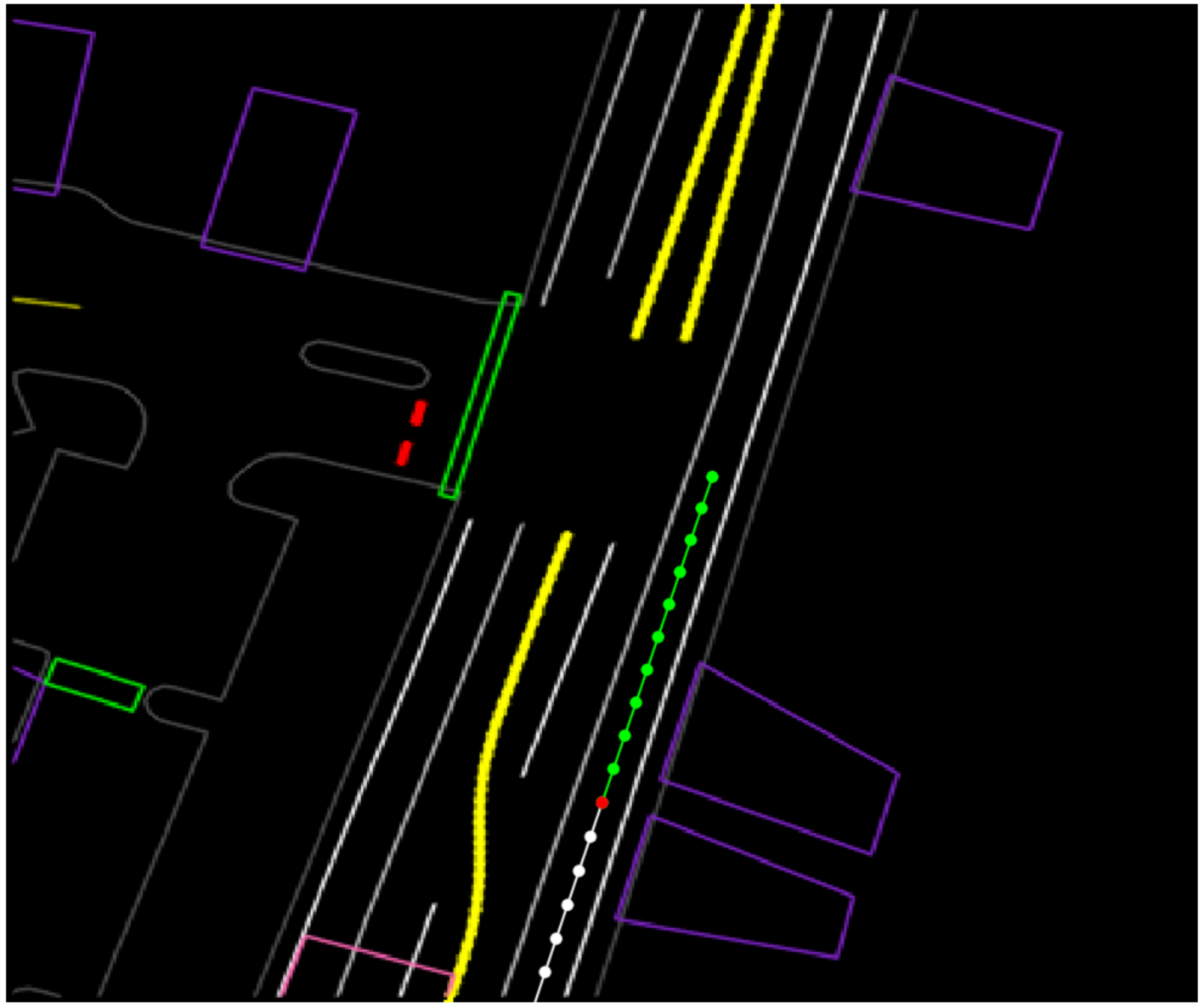
- Векторные фичи – видим дальше
- Трансформер vs CNN – масштабируемость



Фундаментальная проблема imitation learning или behavior cloning



Аугментации



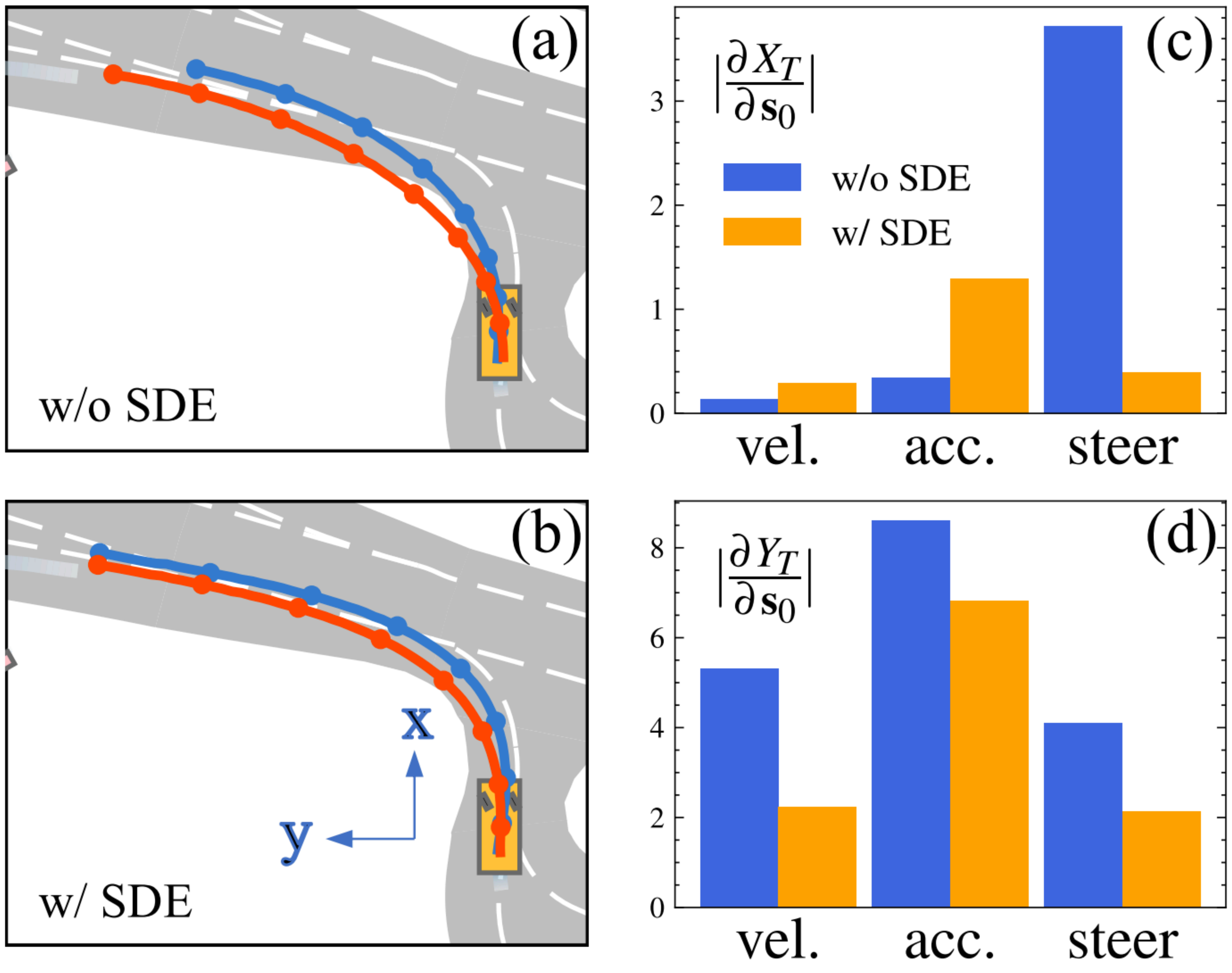
(a) Original

(b) Perturbed

ChauffeurNet

1. Вспомогательные задачи помогают модели лучше обобщаться
    a. Предсказание положений агентов
    b. Предсказания статических препятствий
2. Кроме Imitation Losses нужны лоссы связанные с коллизиями, вылетом с дороги и т.п.
3. Нужны пертрубации траектории

От чего зависит наша траектория?



Еще больше аугментаций

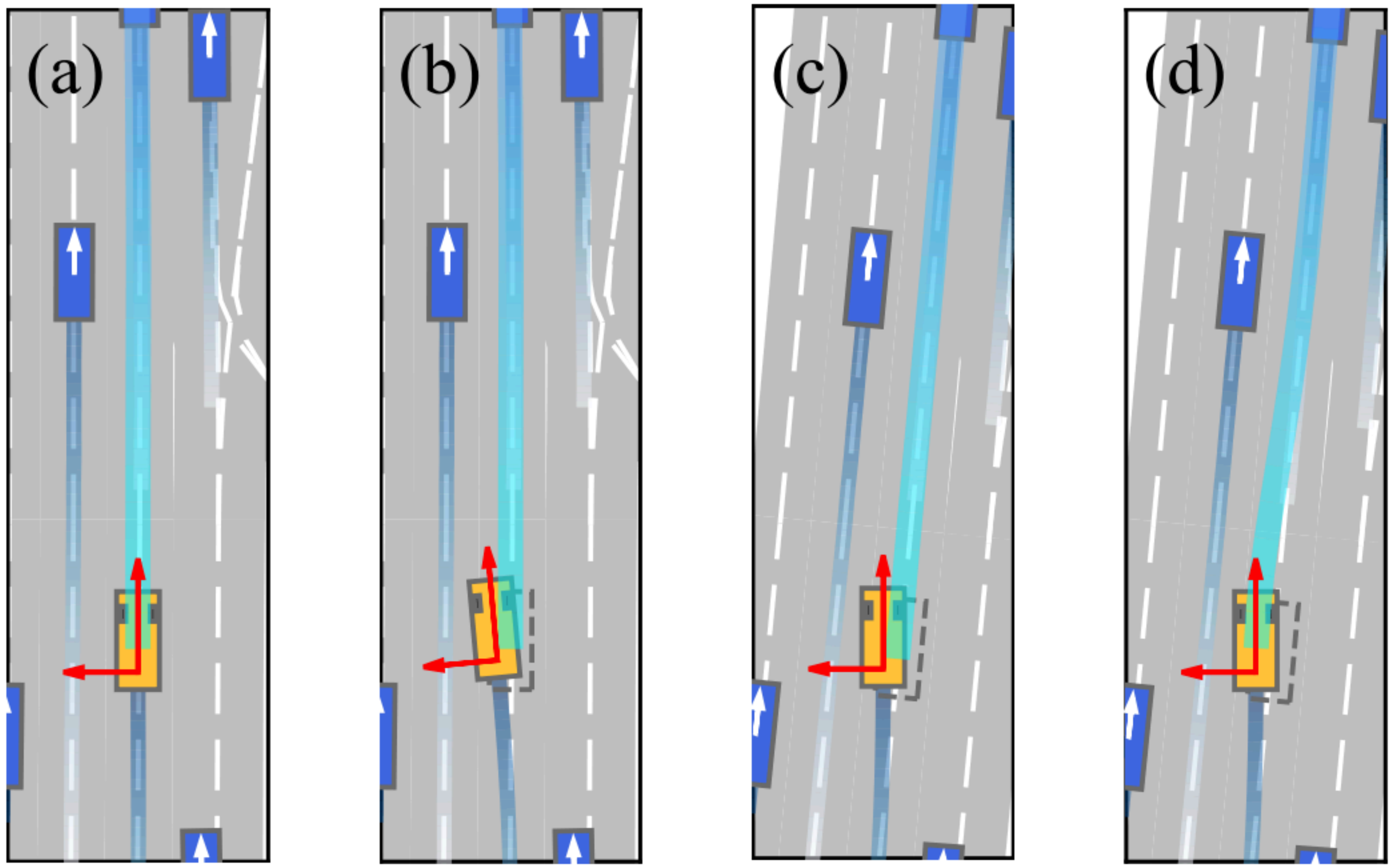


Fig. 4. (a) The original scenario. (b) Random noise is added to the AV's current state and history motion is smoothed. (c) The coordinates of the scenario are re-normalized based on the perturbed position of the AV. (d) A corrected future trajectory is generated using constrained nonlinear optimization.

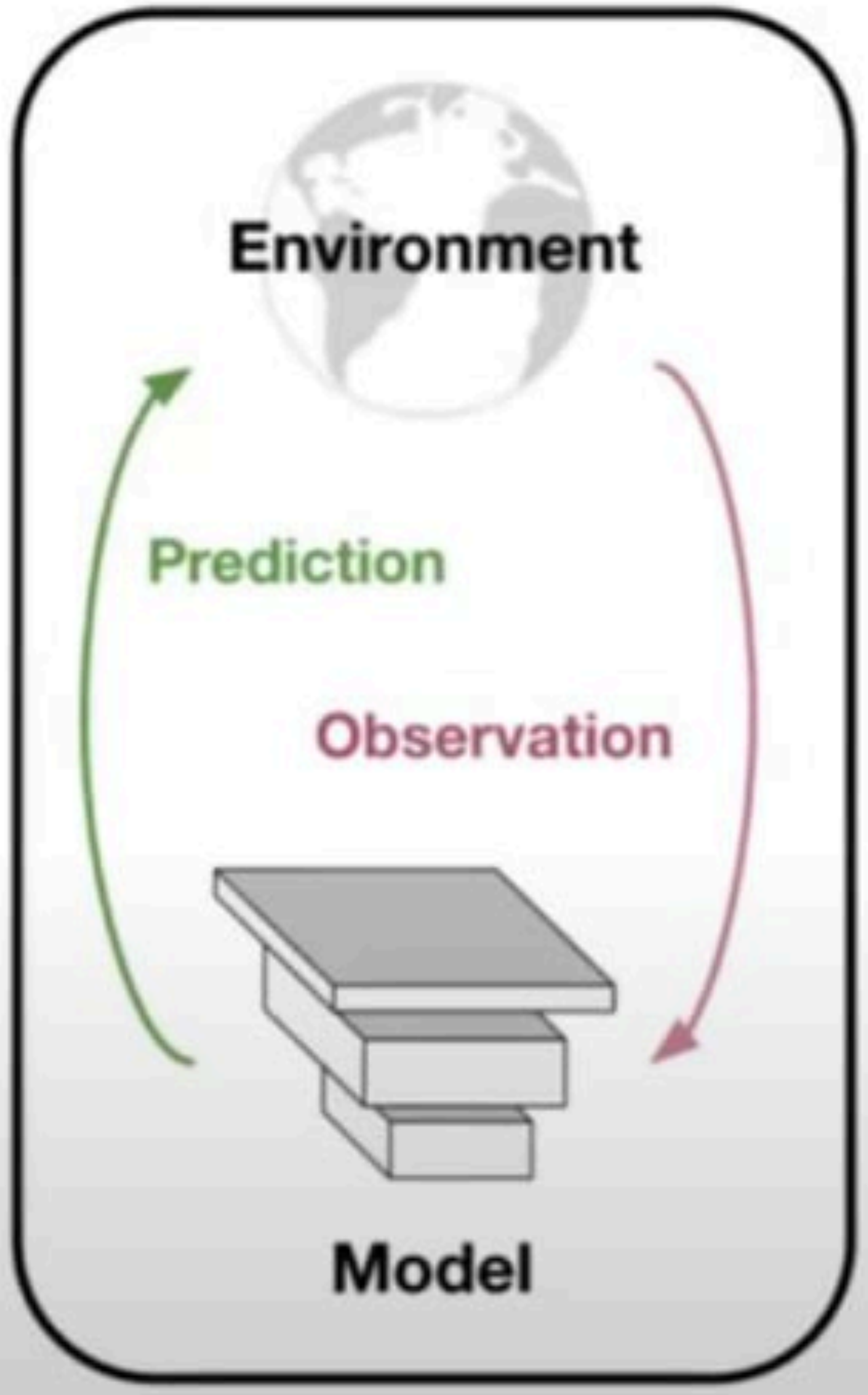
Open loop


Сенсоры
Локализация
Восприятие
Предсказания
Планирование
Управление

Closed loop


Сенсоры
Локализация
Восприятие
Предсказания
Планирование
Управление

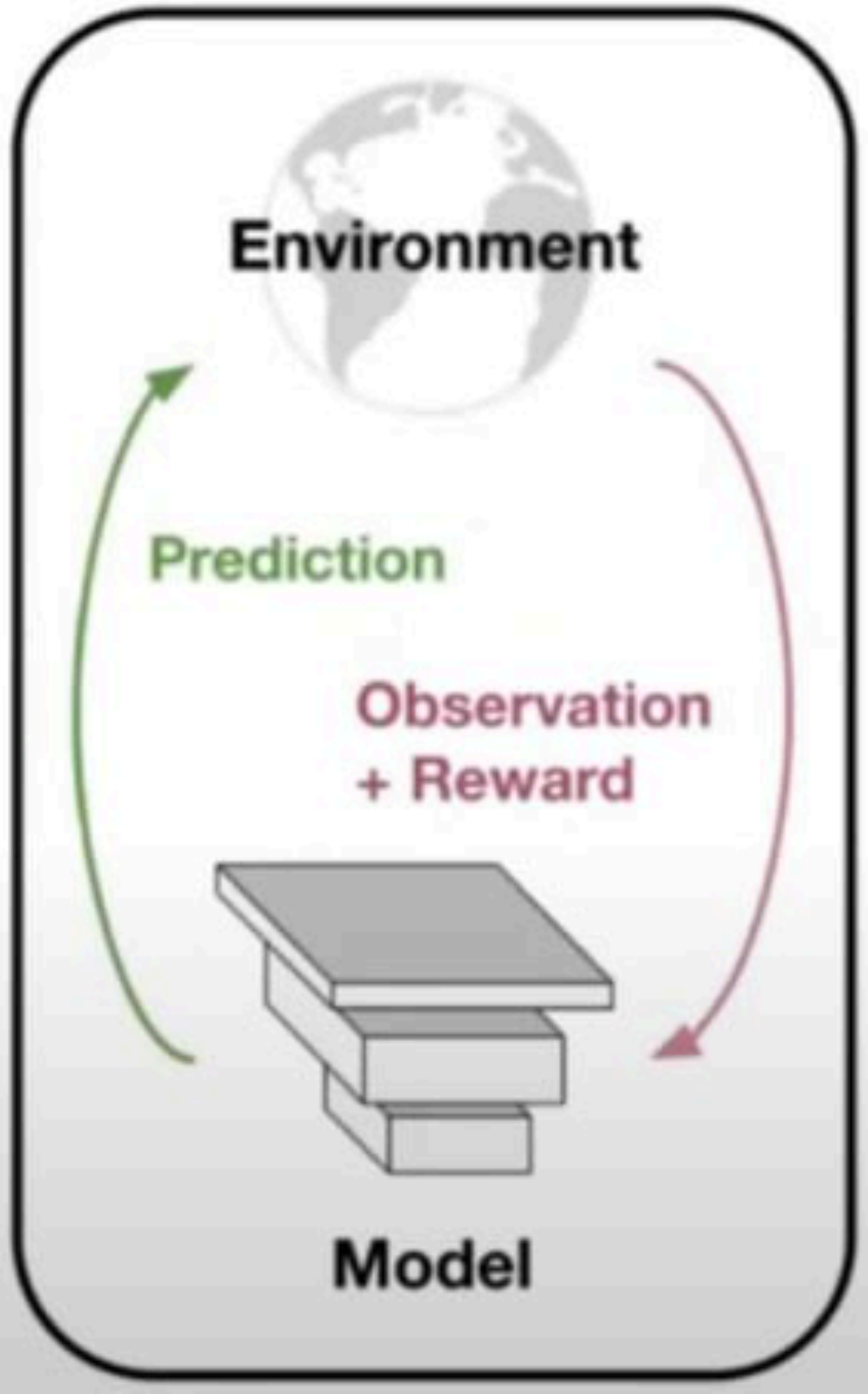
Closed-loop learning



Closed-loop learning:
- Execute policy's actions and see new observations
- No distribution shift between training and inference
- Requires learning in the real world or realistic simulation

Question:
- What is the supervision for new states visited by the model?

Reinforcement learning



Encode explicit learning signal through reward design

RL in the real world[1, 2]

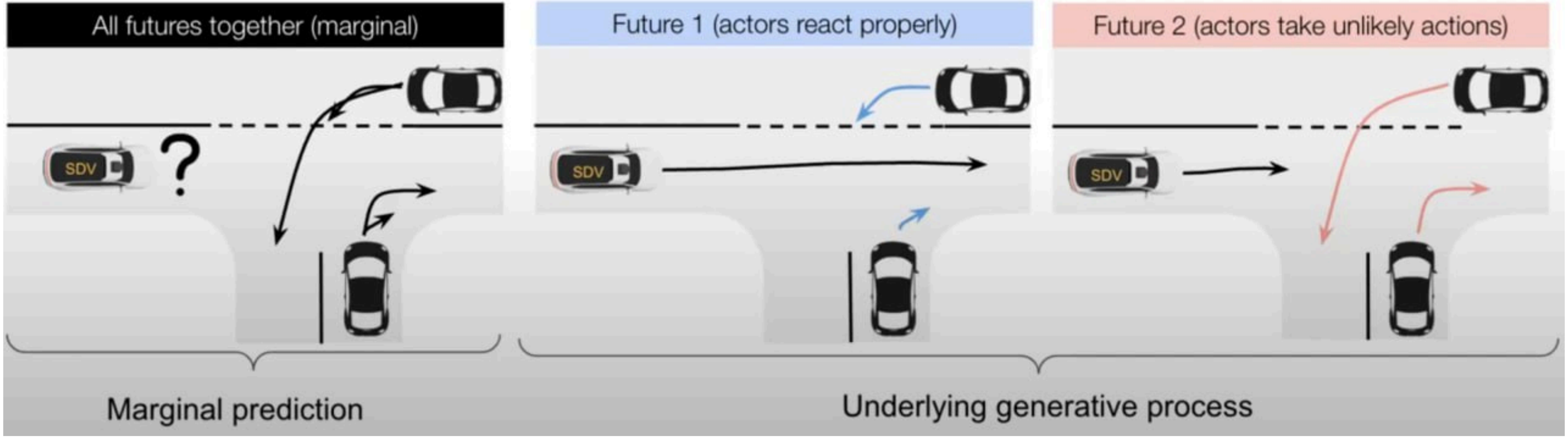
- Eliminates domain gap, but difficult to scale safely

RL in simulation[3]

- Realistic simulation is important to minimize domain gap
- Sample efficiency can be prohibitively low even in simulation

Challenges of planning in an uncertain world

- Challenge: Actors' behaviors can vary based on ego-car's planning
- Challenge: Actors may take very unlikely behaviors in some situations



Q&A