OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

自动驾驶预测规划的联合方法与挑战

刘浩晨

Outline

- 预备知识 / Prerequisites
  - 深度学习及模型基础
  - 线性代数，概率论基本知识
  - 车辆运动学及坐标转换
- 课程目标 / Objectives
  - 掌握预测和规划的基本概念
  - 了解预测规划的通用方法论
  - 熟知评测指标，通用基准与数据集

Outline

- First 10 min
- Second 30min
- Last 20 min

- 预测与规划：背景/ Background
- 预测 / Prediction
  - 预测任务方法论 / Prediction Methodology
  - 预测指标&基准 / Metrics & Benchmarks
- 规划 / Planning
  - 规划框架 / Planning Pipeline
  - 规划指标&基准 / Metrics & Benchmarks
- 课后思考 / Open Questions

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

1. Prediction & Planning: Background / 背景

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

什么是预测&规划？

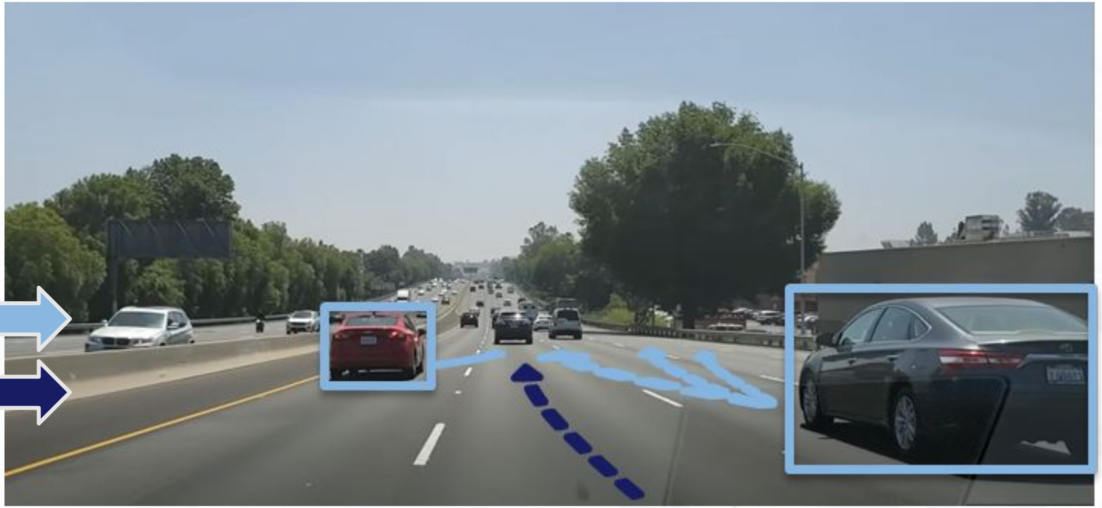
WHATS prediction & planning?

1.1 什么是预测&规划？

WHATS prediction & planning?

憨豆的“自动”驾驶车如是说 [1]:

Mr. Bean’s driver”less” vehicle



预测： 对自车**周边关键交通参与者**(agents)未来状态的估计(estimate)过程

规划： 对自车**从当前状态到完成分段目标**(goal)未来状态的优化(optimize)过程

这里“**规划**”为*局部规划*，而不是*全局规划(地图导航)*

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

为什么需要预测&规划？

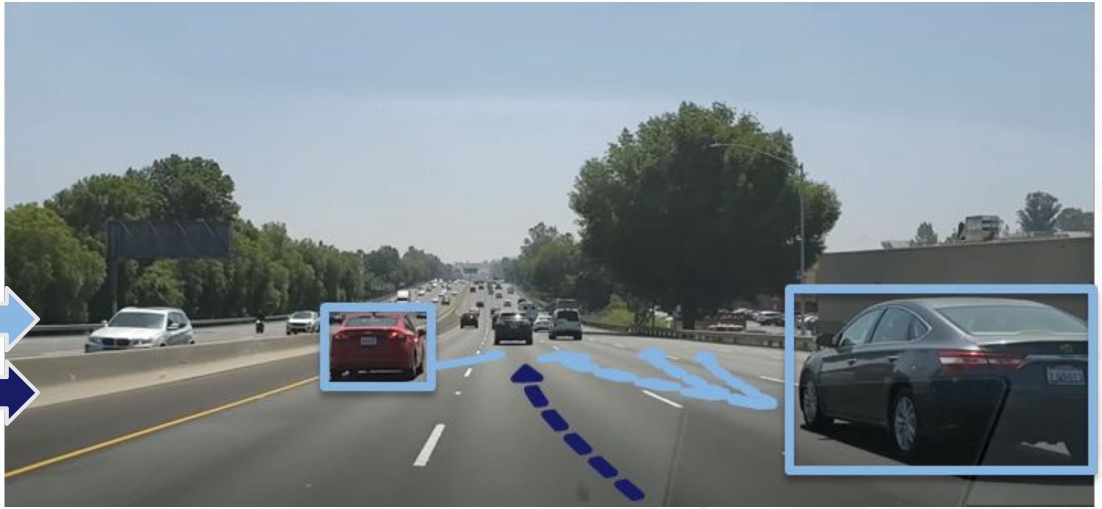
WHY prediction & planning?

1.1 为什么要预测&规划？

WHY prediction & planning?

憨豆的“自动”驾驶车解释道 [1]:

Mr. Bean’s driver”less” vehicle



预测: 为了获得周边参与者未来准确的状态，来指导未来**规划**安全、合规、一致的驾驶行为；

规划: 为了获得自车未来的最优状态，利用**预测**信息进行指引，实现对于未来环境的状态反馈；

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

预测&规划承担什么角色？

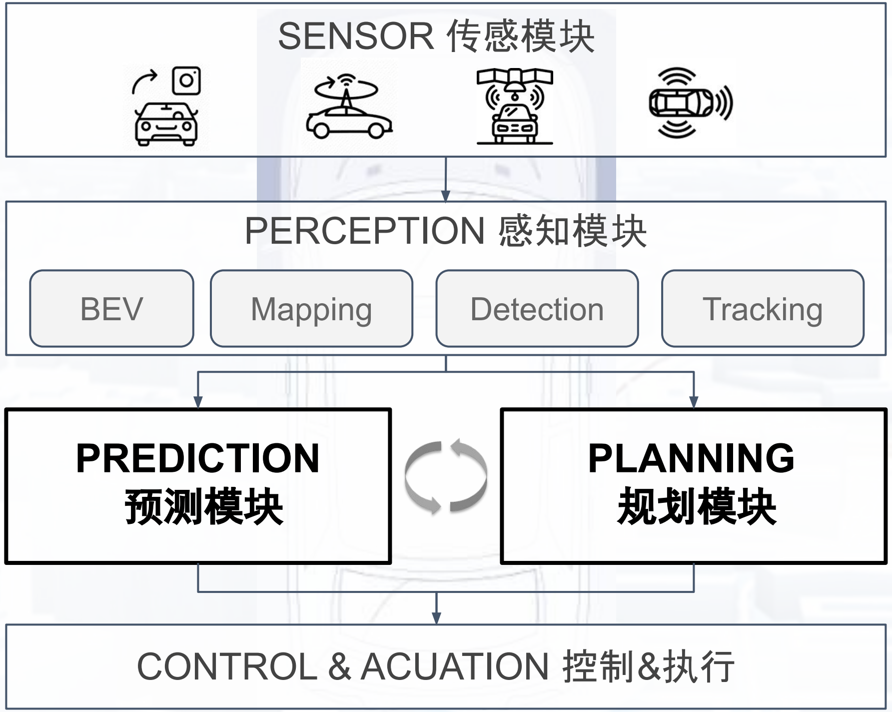
WHERES prediction & planning?

1.2 预测&规划的角色？

WHERES prediction & planning?

憨豆的“自动”驾驶车继续说：

Mr. Bean’s driver”less” vehicle

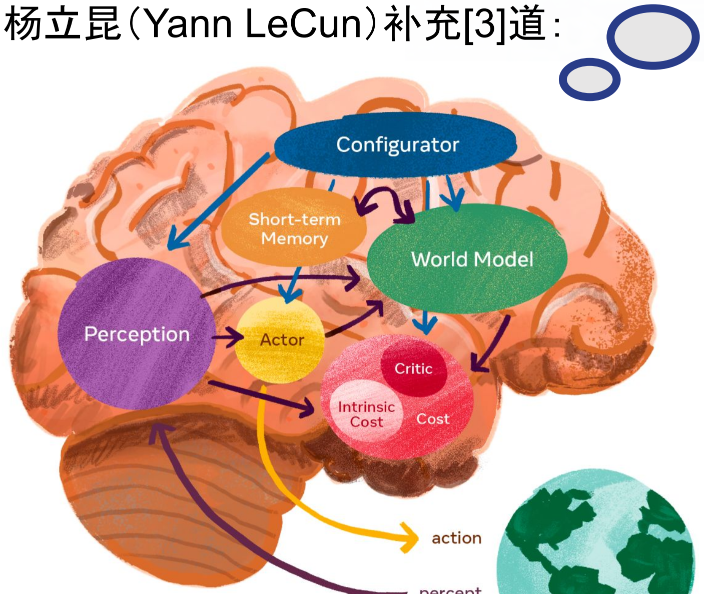


自驾系统中：基于感知(perception)模块输入，输出规划状态予控制(control)执行 [2]

1.2 预测&规划的角色？

WHERES prediction & planning?

杨立昆(Yann LeCun)补充[3]道:



:

预测: 基于感知(Perception),任务(Configurator), 短时记忆(Memory), 和执行策略(Actor)建立对**周边agent的世界模型(World Model)**

规划: 基于感知(Perception),任务(Configurator), 及世界模型的预测(World Model)建立未来状态的**执行策略(Actor)**及**状态评价(Cost)**

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

构建预测&规划框架

HOW do we pipeline prediction & planning?

1.3 如何构建预测&规划框架？

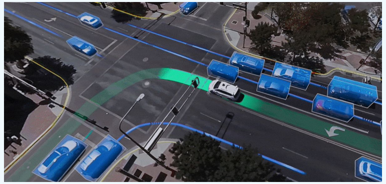
How do we build the **PIPELINE**?

感知

1.任务设定

- 预测？规划？意图OR轨迹预测？(sec2.2,3.2)
- 什么数据集/模拟器？（sec2.4, 3.4)

2.输入信息范式

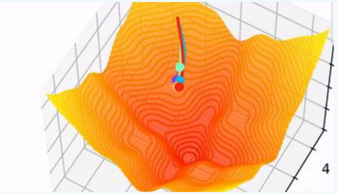


- 确定输入表征(sec 1.4)
- 输入预处理(sec 1.4)

3.模型主干

- 主干设计(sec 1.5)
- 任务模块结合&输出 (sec2.2，3.2)

4.模型学习&优化



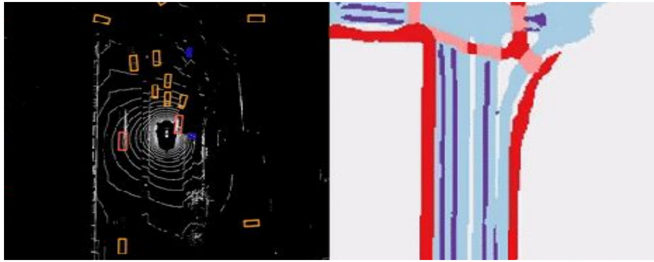
- 确定损失&代价函数
- 模型学习&优化步骤

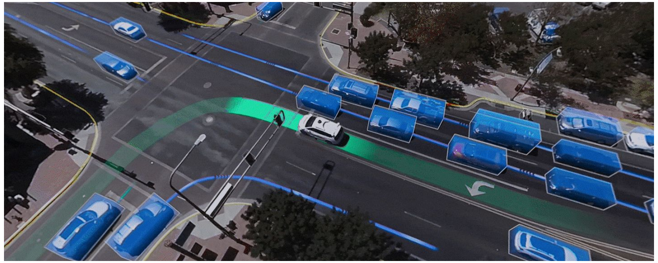
控制

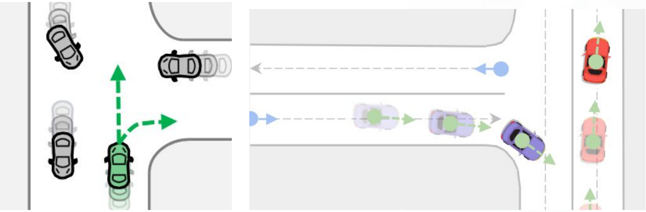
5.基准评测&交互(sec 2.3，3.3)

1.4 我们需要什么输入信息？

What **INPUTS** do we require?







按输入类别[4]:

- 参与者历史状态（Agents）
- 地图信息（Map）

按表征模式[5]:

- 密集/图像化（Dense/Visual）
- 稀疏/向量化（Sparse/Vectorized）

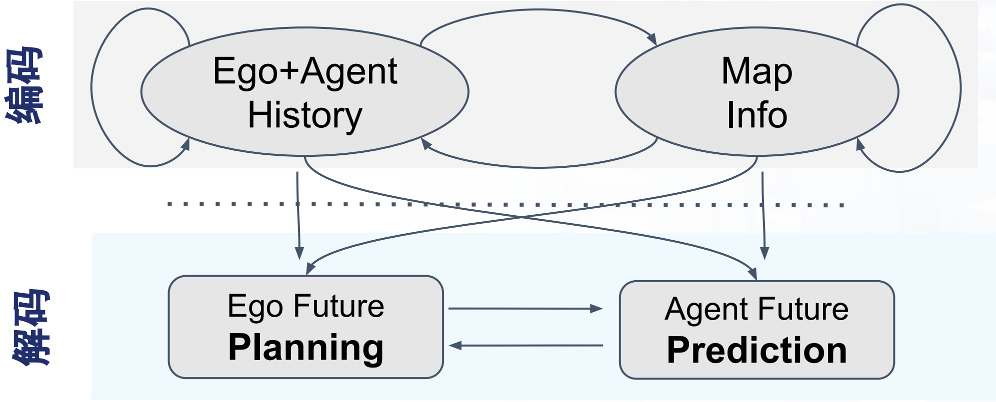
按建立坐标[6]:

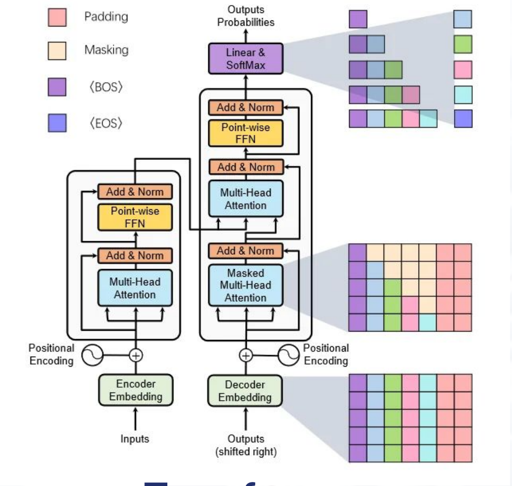
- 自车中心（Ego Centric）
- 对象中心（Object/Query Centric）

1.5 模型主干的核心逻辑？

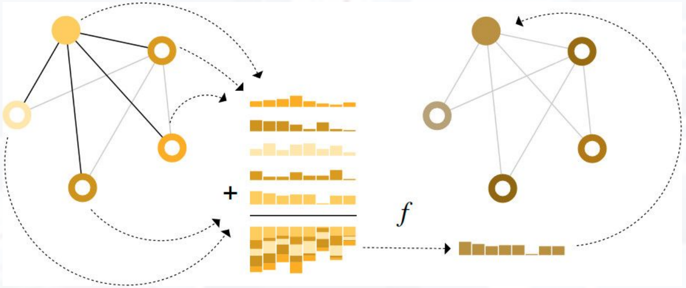
HOW do we model pred. & plan.?

核心: 编码-解码；建模表征间的交互(Interactions)[7]





Transformer



GNN

通用主干(Backbones):

- Transformer (Attention graph) [8]
- 图神经网络 (GNN)[9]

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

2.预测 /Prediction

Outline

Prediction

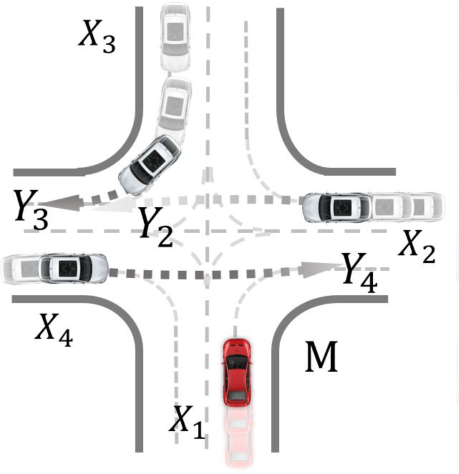
- **定义与挑战** / Definition & Challenges
- **预测任务方法论** / Prediction Methodology
  - **意图预测** / Intention Prediction
  - **单轨迹预测** / Uni-modal Motion Prediction
  - **占用预测** / Occupancy Prediction
  - **多模态预测** / Multi-modal Motion Prediction
- **指标和基准** / Metric & Benchmarks
  - **预测评价指标** / Prediction Metrics
  - **预测基准&数据集** / Benchmarks & Datasets

OpenDriveLab

2.1 定义与挑战

DEFINITIONS & CHALLENGES

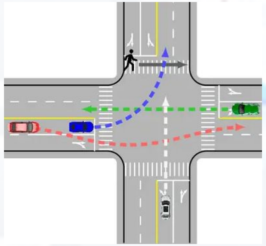
预测： 给定1）自车 X_0及**周边关键交通参与者**(agents) X的历史状态，2）场景地图信息 M；对所有 agents未来状态的估计(estimate)过程[10]： $P(Y_{1:N}|X;M)$



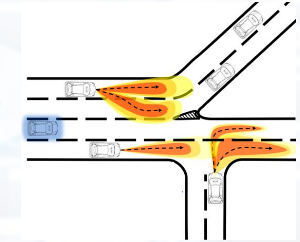
挑战：

- 混杂(heterogenous)的agent类别
- 交互与驾驶场景合规性(compliance)
- 多agents的行为不确定性(uncertainty)

Heterogeneity



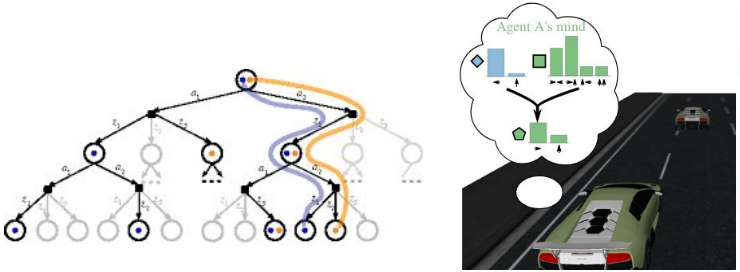
Compliance



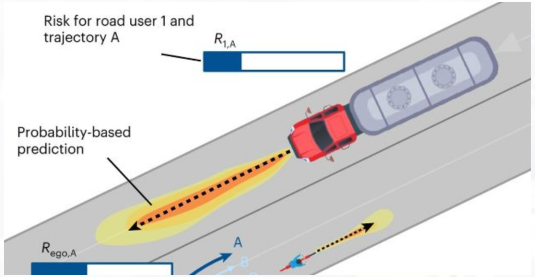
Uncertainty

2.2 预测任务方法论

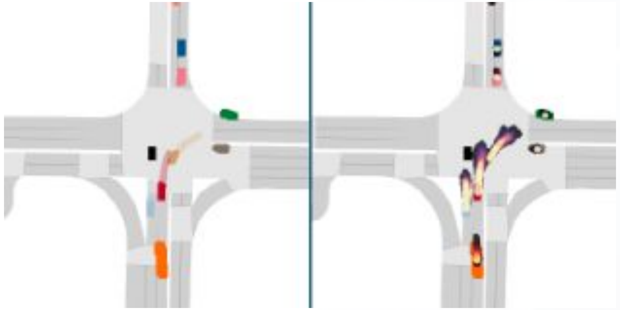
PREDICTION methodology



意图预测(Intention Prediction)



单轨迹预测(Uni-modal Motion Prediction)



占用预测(Occupancy Prediction)

多模态预测(Multi-modal Motion Prediction)

2.2.1 意图预测

Intention Prediction

定义:对于周边agents未来行为类别(behavioral categories)的估计[11]

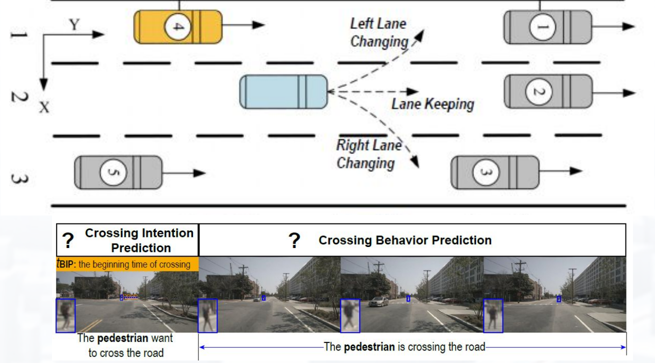
- **分类(classification)任务**

意图输出建模:

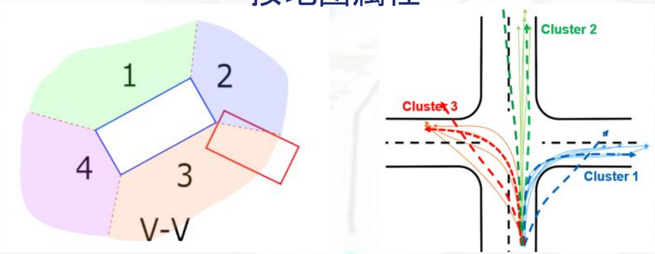
- 按地图属性:(车)换道;(行人)过道 [12, 13]
- 按模式属性:(角度)方位;轨迹模式 [14, 15]
- 按占用属性: 碰撞;交互 [16, 17]

优势:直观可解释的表征,易于建模

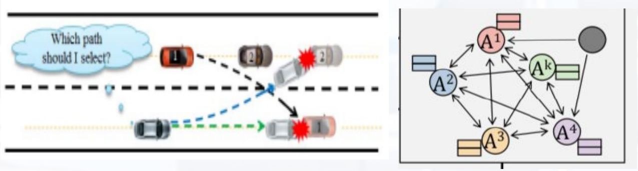
劣势:表征全面性差,模块融合难



按地图属性



按模式属性



按占用属性

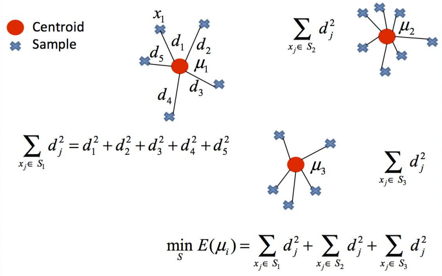
2.2.1 意图预测
Intention Prediction

传统方法:

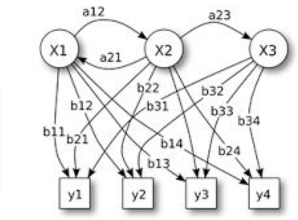
- 场景特征/轨迹聚类[19]
- 概率建模（HMM / Bayes Network）[20, 21]

基于学习的方法:

- 模型：交互关系[22]
- 目标函数：交叉熵（CE/ BCE Focal loss）[23]



KMeans聚类



$$P(\xi_{1:t}|\lambda_i) = \sum_{i=1}^{N} \alpha_t(i)$$

$$\alpha_{t+1}(j) = \left(\sum_{i=1}^{N} \alpha_t(i) \cdot a_{ij}\right) b_j(\xi_{t+1})$$

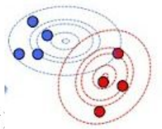
$$\alpha_1(j) = \pi_j b_j(\xi_1)$$

隐马尔可夫模型(HMM)

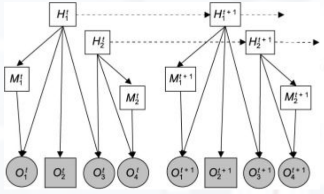
$$\hat{\omega}_k^{j+1} = \frac{1}{n}\sum_{t=1}^{n} P_k^{j+1}(\xi_t)$$

$$\hat{\mu}_k^{j+1} = \frac{\sum_{t=1}^{n}(\xi_t \cdot P_k^{j+1}(\xi_t))}{\sum_{t=1}^{n} P_k^{j+1}(\xi_t)}$$

$$\hat{\Sigma}_k^{j+1} = \frac{\sum_{t=1}^{n}\left(P_k^{j+1}(\xi_t)(\xi_t - \hat{\mu}_k^{j+1})(\xi_t - \hat{\mu}_k^{j+1}\right.}{\sum_{t=1}^{n} P_k^{j+1}(\xi_t)}$$



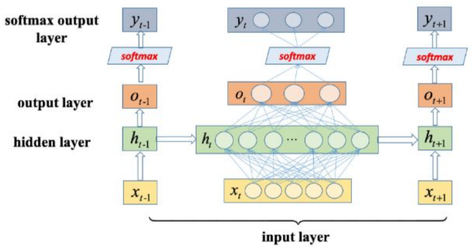
EM聚类



$$r^t(s^t) = \frac{\alpha^t(s^t)\beta^t(s^t)}{\sum_{s^t} \alpha^t(s^t)\beta^t(s^t)}$$

$$S^t = \arg\max_{s^t}(r^t(s^t))$$

贝叶斯网络(Bayes Network)



$$FL = -(1 - P_t)^{\gamma} \log(P_t)$$

$$CE = -\log(P_t)$$

基于学习的方法

2.2.2 单轨迹预测

Uni-modal Motion Prediction

定义:对于周边agents未来状态/轨迹(motion/states)的估计 [24]

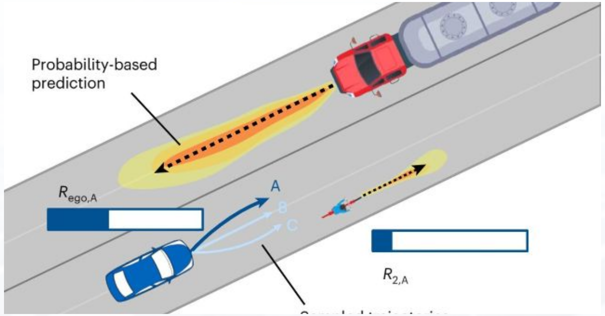
- **回归(regression)任务**

轨迹输出建模:

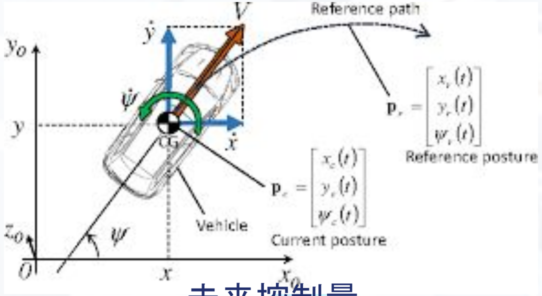
- 未来状态 (x, y, vx, vy, θ)[25]
- 未来控制量 (acc, steer)[26]
- 未来控制点 (bezier, poly-curve)[27]

优势:便于模块间融合，表征统一

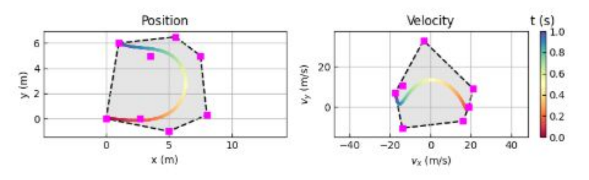
劣势:不确定性/交互影响大，难以学习



未来状态



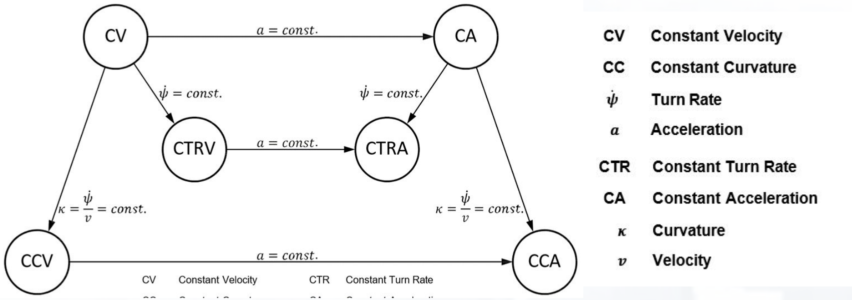
未来控制量

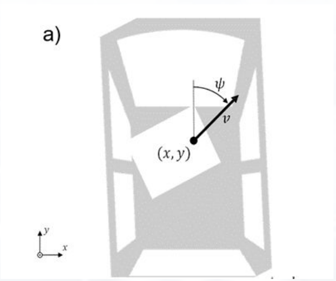


未来控制点

2.2.2 单轨迹预测

Uni-modal Motion Prediction

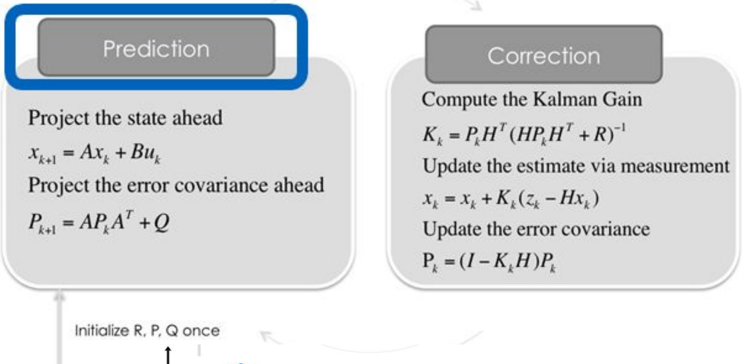


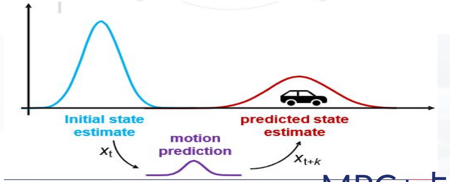


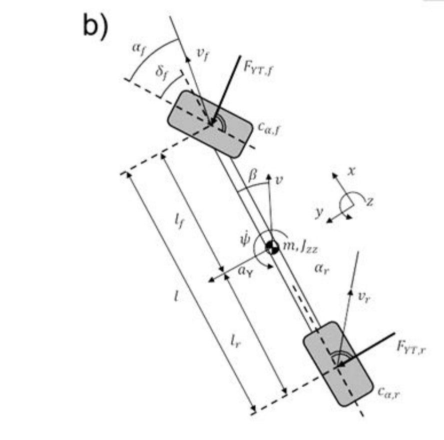
传统方法:

- 物理预测(质点模型)[28]
- 模型预测控制(MPC)+卡尔曼滤波器(Kalman Filter)(运动学模型)[29]

物理预测(质点模型)







MPC+卡尔曼:预测-校正

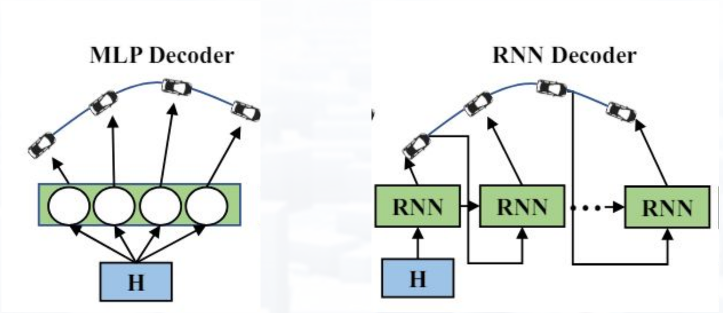
2.2.2 单轨迹预测

Uni-modal Motion Prediction

基于学习方法:

模型解码方法[30]:

- 直接解码（MLP-oneshot）
- 自回归解码（RNN-autoregressive）



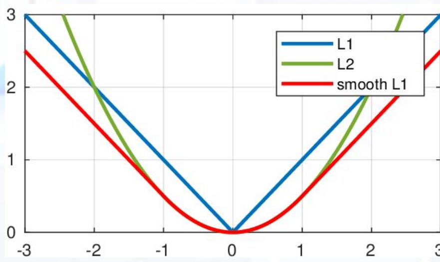
模型解码分类

模型更新方法[31, 32]:

- 基于回归的方法（L1）
- 基于概率的方法(Negative Log-Likelihood + Gaussian & Laplace）

$$l_n = \begin{cases} 0.5(x_n - y_n)^2/beta, & \text{if } |x_n - y_n| < beta \\ |x_n - y_n| - 0.5 * beta, & \text{otherwise} \end{cases}$$

Smooth L1



$$-\log \mathcal{N}_h(\hat{Y}_x - \mu_x, \sigma_x; \hat{Y}_y - \mu_y, \sigma_y; \rho)$$

NLL-Gaussian

$$-\frac{1}{NH}\sum_{i=1}^{N}\sum_{t=T+1}^{T+H} \log P\left(\mathbf{R}_i^\top\left(\mathbf{p}_i^t - \mathbf{p}_i^T\right) \mid \hat{\boldsymbol{\mu}}_i^t, \hat{\mathbf{b}}_i^t\right)$$

NLL-Laplace

2.2.3 占用预测

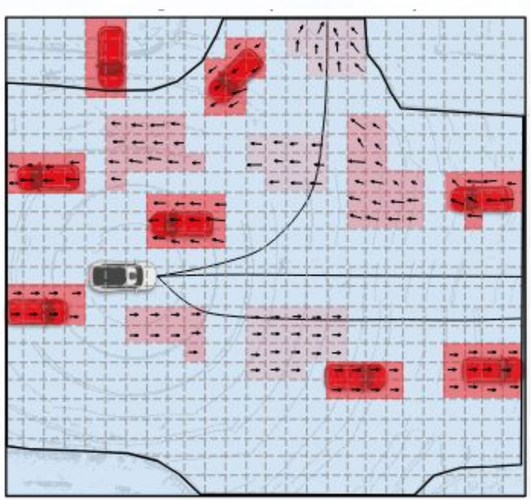
Occupancy Prediction

定义:对于周边密集空间单位，未来是否存在agent占据的估计 [10]

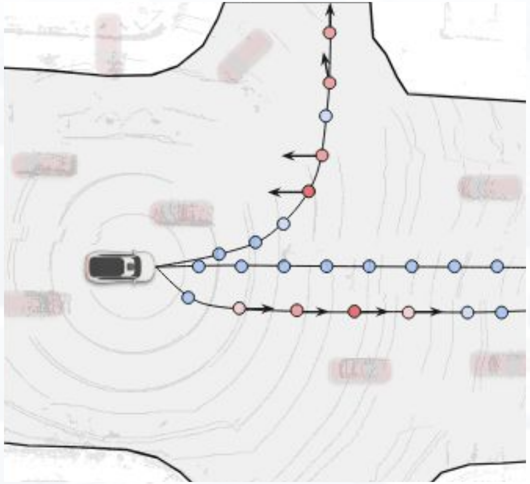
- **二分类(binary classification)任务**

占用预测建模:

- 鸟瞰(BEV) 占用预测 [43]
- 隐式(implicit) 占用预测 [44]



BEV占用预测



隐式占用预测

优势:**BEV感知对齐，预测任意agents数量**

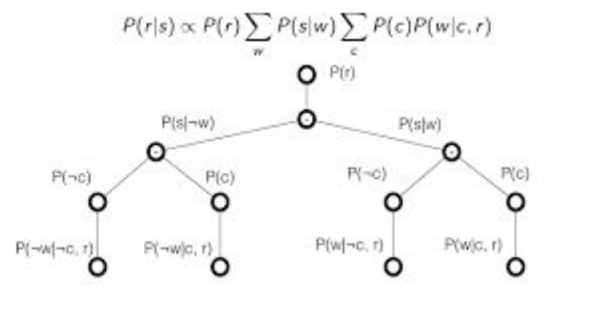
劣势:**难以跟踪(tracking)，计算开销大**

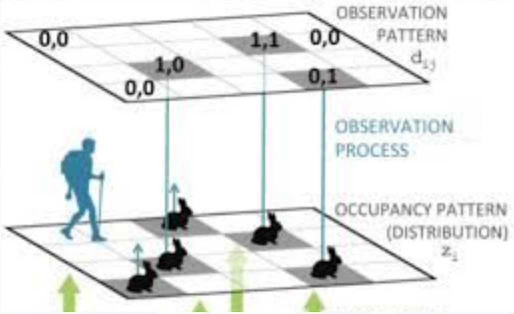
2.2.3 占用预测

Occupancy Prediction

传统方法:

- 基于贝叶斯方法[45]



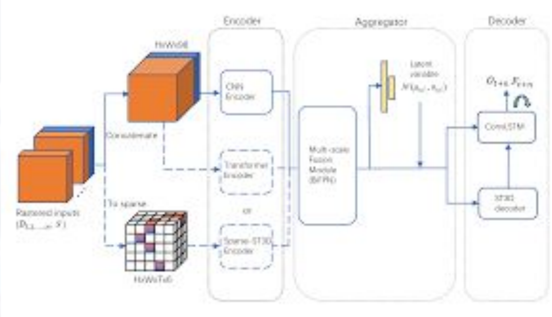


基于学习的方法:

- 目标函数:二元交叉熵(BCE/Focal loss)[43]

$$FL = -(1 - P_t)^{\gamma}\log(P_t)$$

$$CE = -\log(P_t)$$



2.2.4 多模态预测

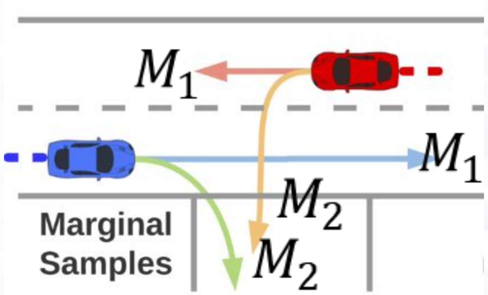
Multi-modal Motion Prediction

定义:对于周边agents未来行为模态和对应轨迹状态的估计[33]

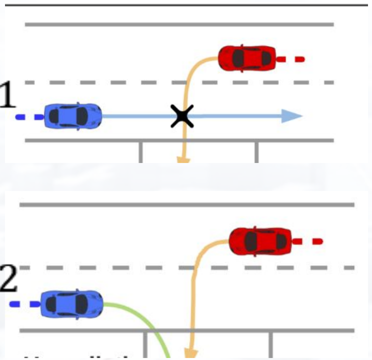
- **模态:分类(classification)任务**
- **轨迹:回归(regression)任务**

多模态预测输出建模:

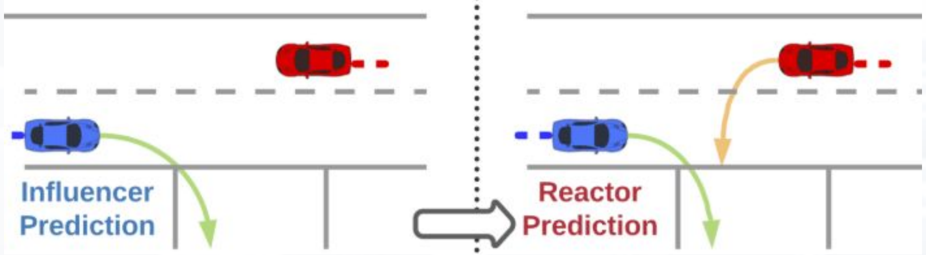
- 边际预测(marginal)[34]
- 联合预测 (joint)[35]
- 条件预测 (conditional)[36]



边际预测



联合预测



条件预测

优势:考虑行为/场景不确定性,类人模式

劣势:模态间冲突,模态崩塌现象

2.2.4 多模态预测

Multi-modal Motion Prediction

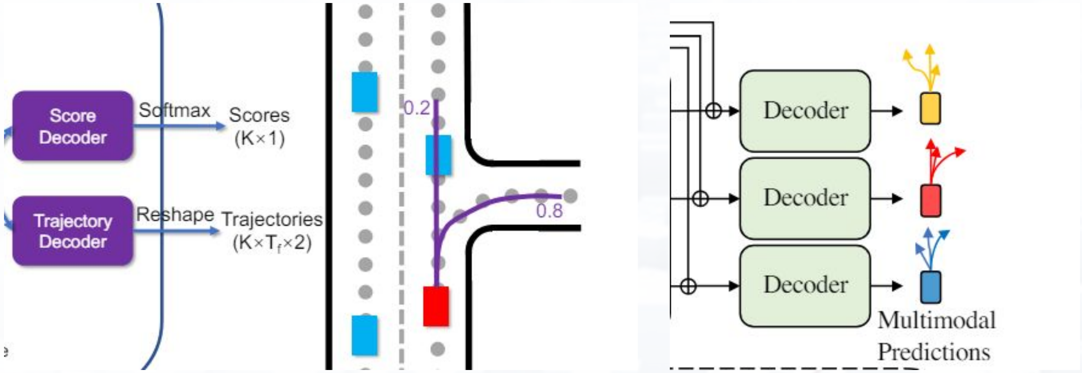
解码生成方法:

直接回归 (Direct Regression)[37]:

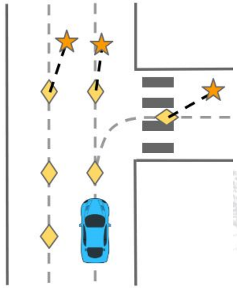
- 打分解码模态概率
- 轨迹解码对应模态的轨迹

先验回归(Prior Regression):

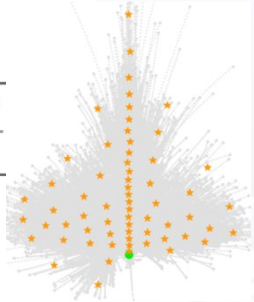
- 目标(goal)先验预测[38]
- 锚点(anchor)先验预测[32]
- 地图(map)先验预测[39]
- 热力图(heatmap)预测[40]
- 轨迹集召回(retrieval)[41]



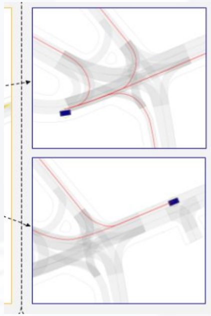
直接回归法



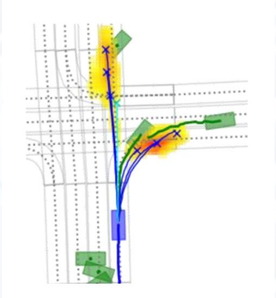
目标先验



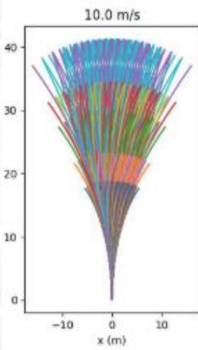
锚点先验



地图先验



热力图先验

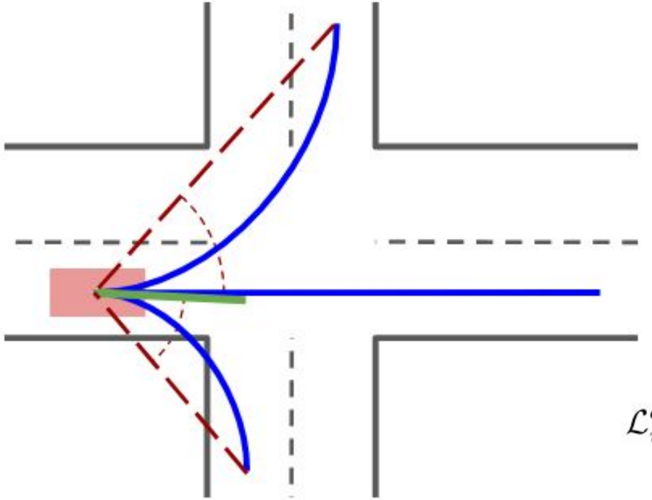


轨迹集召回

2.2.4 多模态预测

Multi-modal Motion Prediction

更新方法:



$$\sum_{m=1}^{M} p_{im} L(\boldsymbol{\tau}_{ij}, \tilde{\boldsymbol{\tau}}_{imj}),$$

$$m^* = \mathop{\arg\min}_{m\in\{1,\ldots,M\}} dist(\boldsymbol{\tau}_{ij}, \tilde{\boldsymbol{\tau}}_{imj}).$$

$$\mathcal{L}_{ij}^{class} = -\sum_{m=1}^{M} I_{m=m^*} \log p_{im},$$

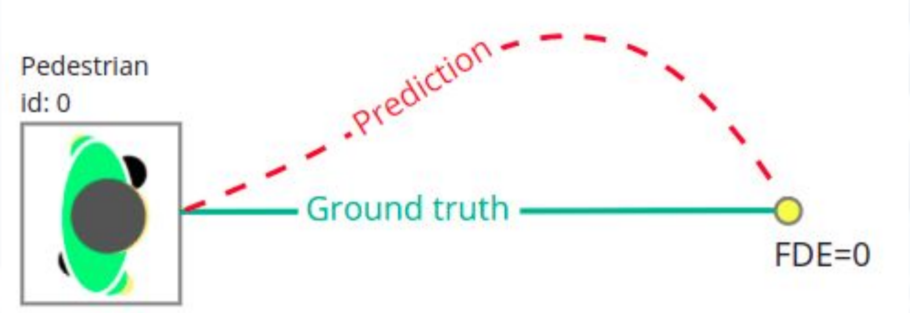
$$\mathcal{L}_{ij}^{class} + \alpha \sum_{m=1}^{M} I_{m=m^*} L(\boldsymbol{\tau}_{ij}, \tilde{\boldsymbol{\tau}}_{imj}),$$

WTA：预测轨迹与GT最接近的模态被更新

模态概率:

- 动态更新："赢者通吃"(Winner take all loss，WTA loss)[42]
- 静态更新：先验相似度确定标签[32, 39, 40]

预测轨迹:



- 平均偏差(ADE)[42]
- 末端偏差(FDE)[37]

参考2.2.2

2.3 预测指标

PREDICTION Metrics

1. 意图预测:

- Acc；Precision；Recall
- F1Score； AUC

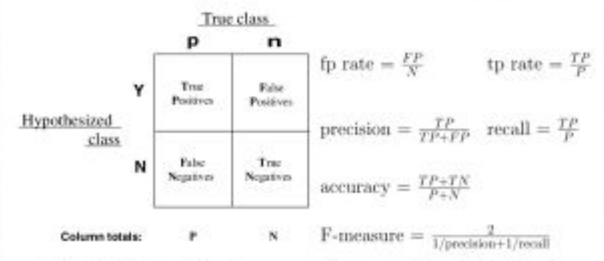
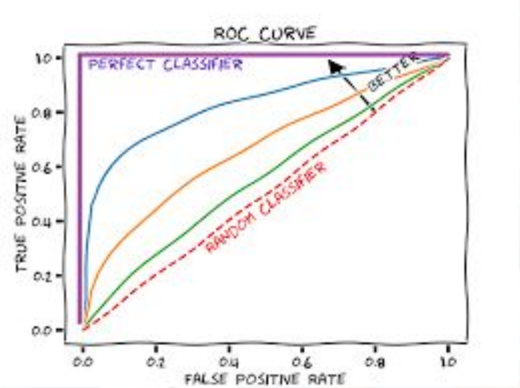


Fig. 1. Confusion matrix and common performance metrics calculated from it.



2. 单轨迹预测:

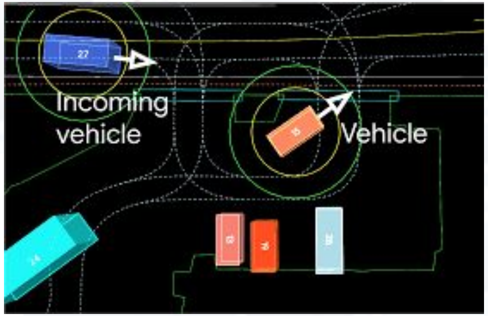
- ADE；FDE
- 遗漏率（Miss Rate）

$$ADE = \frac{1}{T}\sum_{t=1}^{T}\sqrt{(x_i - x_i^{GT})^2 + (y_i - y_i^{GT})^2}$$

$$FDE = \sqrt{(x_T - x_T^{GT})^2 + (y_T - y_T^{GT})^2}$$

$$f(.) = \mathbb{1}[x_a^k > \lambda^{lon}] \vee \mathbb{1}[y_a^k > \lambda^{lat}]$$

$$[x_a^k, y_a^k] := (\hat{s}_a - s_a^k) \cdot \mathbf{R}_a$$

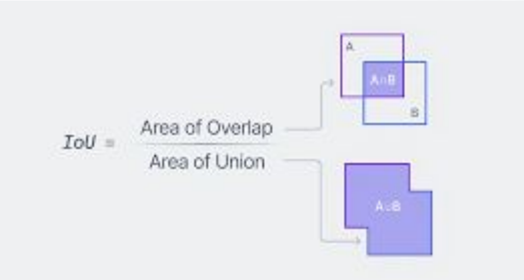


2.3 预测指标

PREDICTION Metrics

3. 占用预测:

- AUC
- 交并比(IoU) / Soft IoU



$$\text{Soft-IoU}(O_t^{\mathcal{K}}, \tilde{O}_t^{\mathcal{K}}) = \frac{\sum_{x,y} O_t^{\mathcal{K}} \cdot \tilde{O}_t^{\mathcal{K}}}{\sum_{x,y} O_t^{\mathcal{K}} + \tilde{O}_t^{\mathcal{K}} - O_t^{\mathcal{K}} \cdot \tilde{O}_t^{\mathcal{K}}}$$

4. 多模态预测:

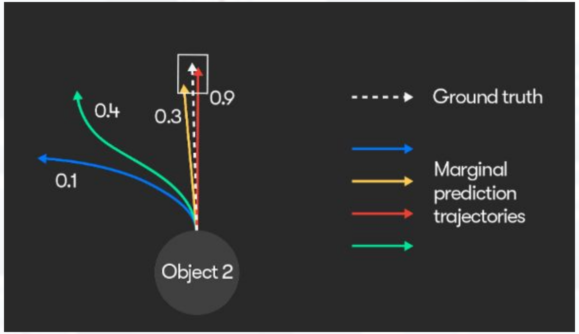
- minADE / minFDE
- BrierminADE / BrierminFDE
- Mean AP (mAP) / Soft mAP

$$\text{minADE}(G) = \min_i \frac{1}{T} \sum_{t=1}^{T} \left\| \hat{s}_G^t - s_G^{it} \right\|_2$$

$$\text{brier-ADE} = \text{ADE} + (1-p)^2,$$

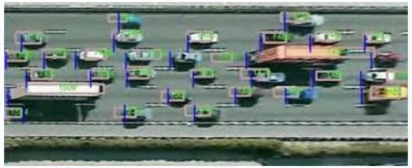
$$\text{brier-FDE} = \text{FDE} + (1-p)^2.$$

$$\text{AP} = \int_0^1 p(r)dr$$



2.4 预测基准&数据集

PREDICTION benchmark & Dataset



公开数据集:

- NGSIM(3x15min，高速公路，标注数据)[46]
- HighD(11.5h，高速公路，标注数据)[47]
- NuScenes (1000场景，全数据)[48]
- INTERACTION (16.5h，4w场景，标注数据)[49]
- Argoverse 1 (320h，20w场景，标注数据)[50]
- Argoverse 2 (763h，25w场景，标注/全数据)[51]
- Waymo Motion (574h，56w场景，标注/全数据)[52]

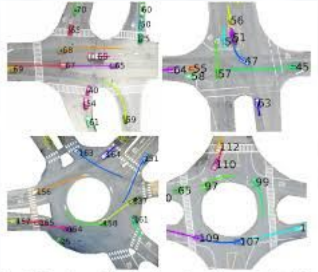


Fig. 1: Examples of the detection and tracking results in highly

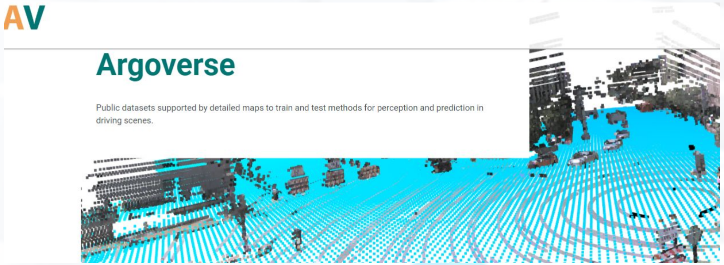
2.4 预测基准&数据集

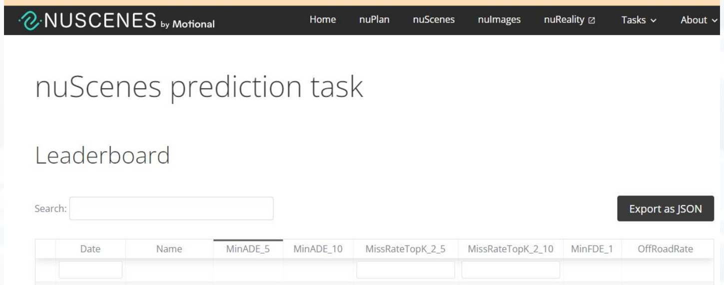
PREDICTION benchmark & Dataset

公开基准/排行榜:

1. 边际(marginal)多模态预测:

- Waymo Motion Challenge（8s, 6 modals, Soft mAP）
- Argoverse 1 (3s, 6 modals, BrierminFDE)
- Agroverse 2 (6s, 6 modals, BrierminFDE)
- INTERACTION (3s, 6 modals, Miss Rate)
- NuScenes (8s, 5 modals, minADE)

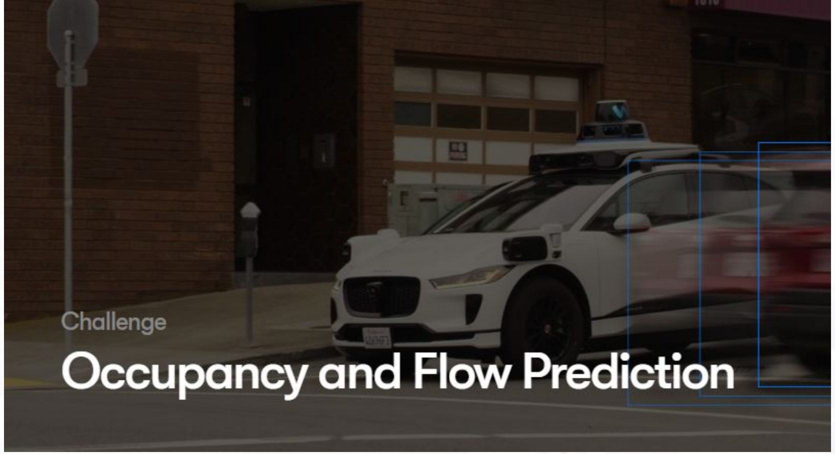




2.4 预测基准&数据集

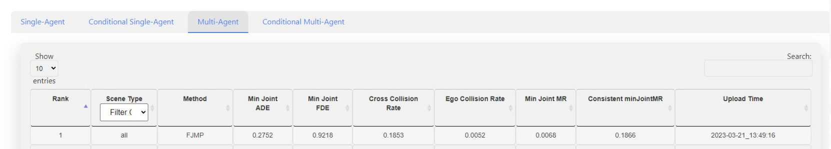
PREDICTION benchmark & Dataset

公开基准/排行榜:



2. 联合(joint)多模态预测:

- **Waymo Interaction Challenge（8s, 6 modals, 2 agents, mAP）**
- **Argoverse 1 (3s, 6 modals, BrierminFDE)**
- Agroverse 2 (6s, 6 modals, BrierminFDE)
- INTERACTION (3s, 6 modals, Miss Rate)



Single-Agent | Conditional Single-Agent | Multi-Agent | Conditional Multi-Agent

Show 10 entries

Search:

Rank	Scene Type	Method	Min Joint ADE	Min Joint FDE	Cross Collision Rate	Ego Collision Rate	Min Joint MR	Consistent minJointMR	Upload Time
1	all	FJMP	0.2752	0.9218	0.1853	0.0052	0.0068	0.1866	2023-03-21_13:49:16

3. 条件(conditional)多模态预测:

- INTERACTION (3s, 6 modals, Miss Rate)

4. 占用预测:

- Waymo Occupancy（8s, AUC)

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

3.规划 /Planning

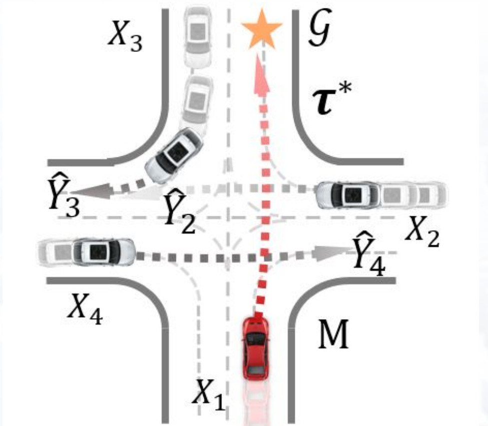
Outline

Planning

- **定义与挑战** / Definition & Challenges
- **规划框架** / Planning Pipeline
  - **候选轨迹生成** / Motion Profiles Generation
  - **规划范式** / Planning Paradigm
  - **代价函数** / Cost function
  - **优化器** / Optimization Tools
- **指标基准** / Metric & Benchmarks

3.1 定义与挑战

DEFINITIONS & CHALLENGES

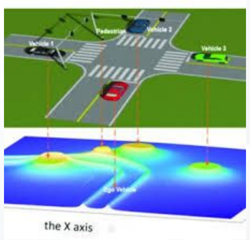


规划: 给定1)自车 X_0 及**周边关键交通参与者**(agents) $X_{1:N}$ 的历史状态，2)场景地图信息 M；3)周边agents的预测 $\hat{Y}_{1:N}$；4)目标信息 $\mathcal{G}$；对自车未来短期状态 $\hat{Y}_0$ 的优化(optimize)[53]

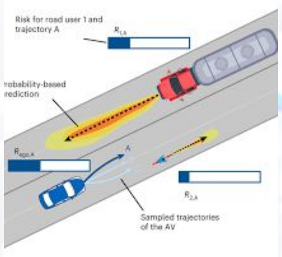
$$\tau^* = \underset{\tau\subset\hat{Y}_0}{\operatorname{argmin}}\, \mathbf{C}(\hat{Y}, X, M, \mathcal{G})$$

挑战:

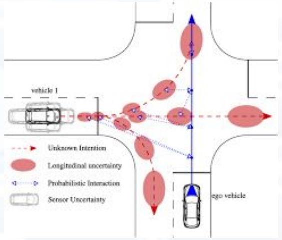
- 优化/代价函数设计鲁棒性(robustness)
- 同预测耦合的合规性(compliance)
- 行为不确定性(uncertainty)



Robustness



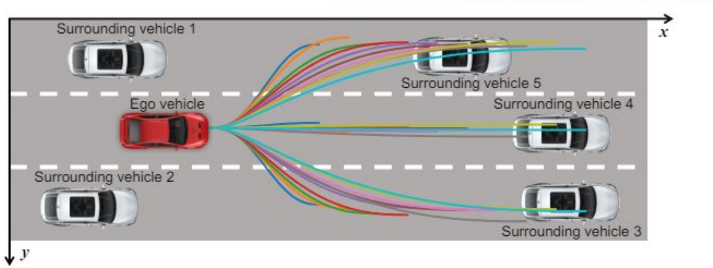
Compliance



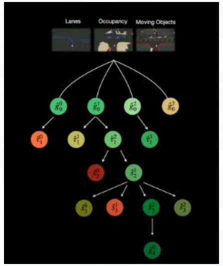
Uncertainty

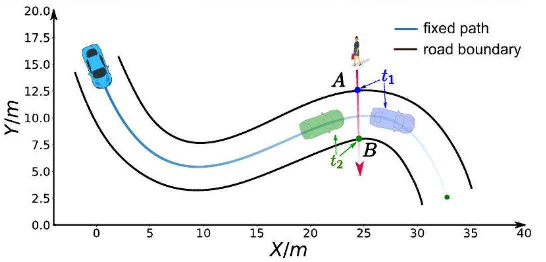
3.2 规划框架

PLANNING pipeline

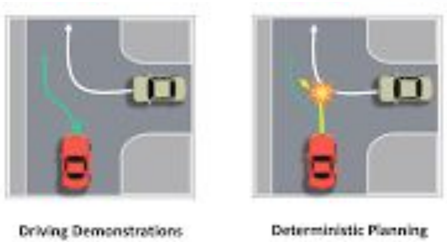


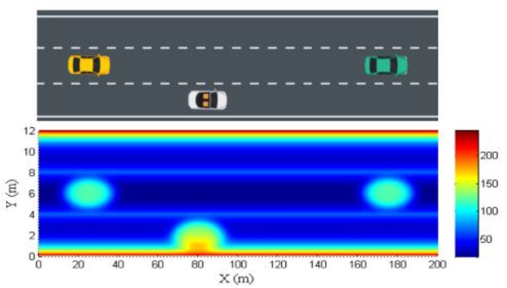
1.候选轨迹生成（Motion Profile Generation）

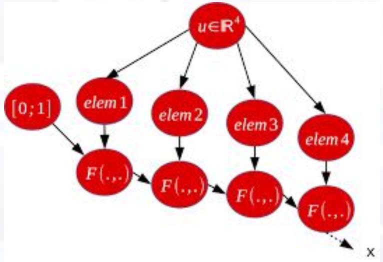




2.规划范式（Planning Paradigm）







3.代价函数与优化器（Cost functions & Optimization Tools）

3.2.1 候选轨迹生成

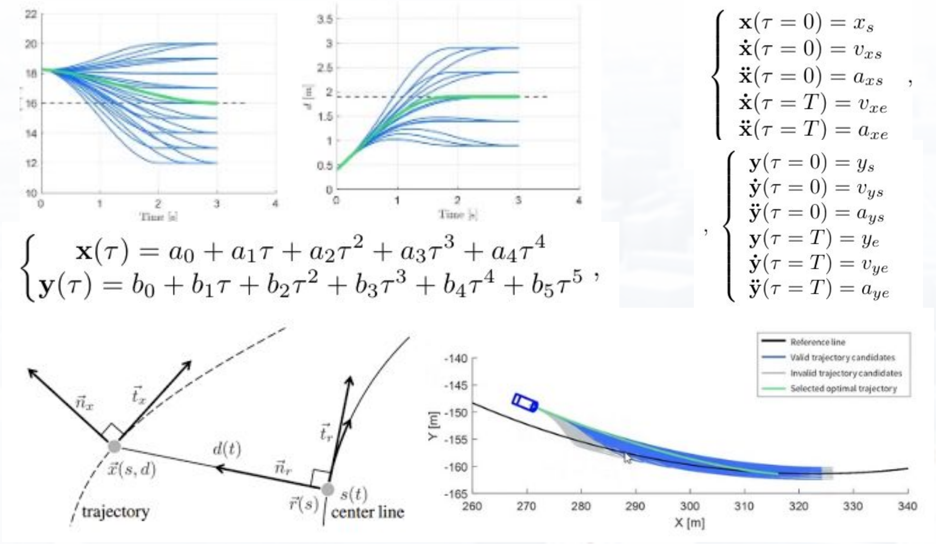
MOTION Profile Generation

多项式轨迹生成[54]:

1. 基于多项式公式和控制条件进行生成
2. 中心线坐标(Frenet)转换[55]

优势:规划轨迹平滑,保证丰富程度

劣势:依赖准确中心线和速度范围



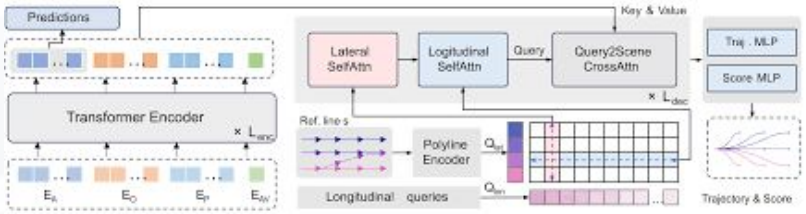
多项式轨迹生成

模仿学习(IL)轨迹生成[56, 57]:

1. 建立交互规划模型(参考1.4)
2. 生成单/多模态规划轨迹(参考2.2.3)

优势:类人规划,易结合,考虑不确定性

劣势:平滑性,丰富度差别大



模仿学习轨迹生成

3.2.2 规划范式
PLANNING Paradigm

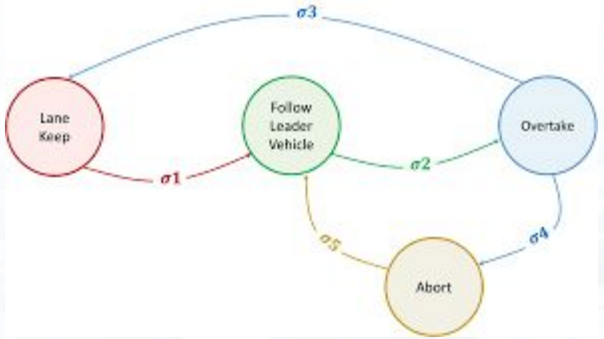
1. **行为规划(behavior planning)**:对自车离散驾驶行为的优化及对应轨迹的链接 [58]

- 基于有限状态机(FSM)的优化 [59]
- 基于逆强化学习(IRL)的优化 [60]
- 基于搜索+动态规划(Search+DP)的优化 [61]

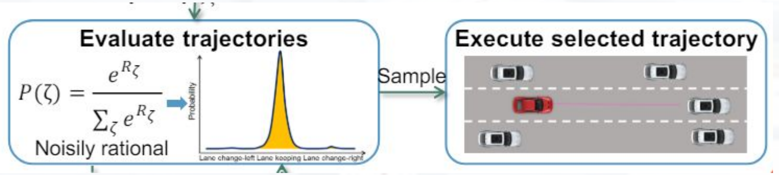
优化步骤

1. 计算单步 / 多步行为的代价/value(cost/value)
2. 选择代价最小的行为组合，连接对应候选轨迹

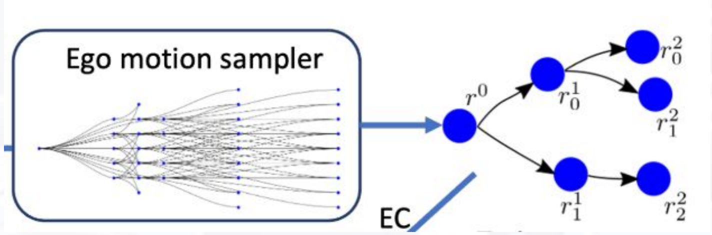
优势:优化速度快，便于模块融合，可解释
劣势:规划连续性弱，规划跳变等



有限状态机(FSM)



逆强化学习(IRL)

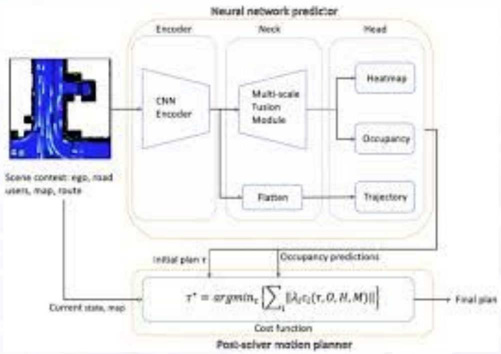


树搜索(Tree Search)

3.2.2 规划范式
PLANNING Paradigm

2. 运动规划（motion planning）：对自车连续状态轨迹的优化 [58]

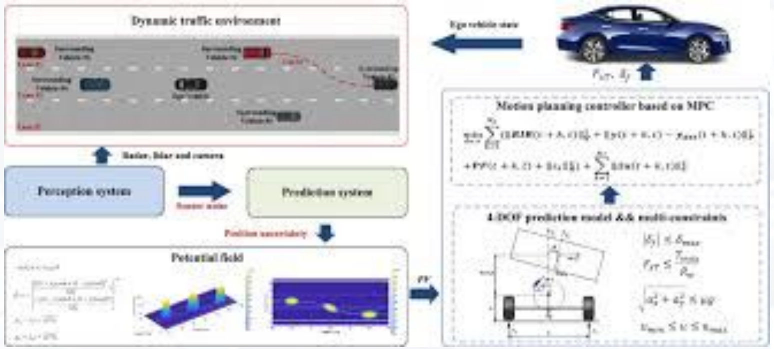
- 无模型（Model-free）轨迹优化 [62]
- 模型预测控制（MPC）的优化 [63]



无模型轨迹优化（Model-free）

优化步骤

1. 给定候选轨迹作为优化变量初值
2. 建模单步 / 多步规划的代价/value（cost/value）
3. 通过优化器对应候选轨迹进行优化



模型预测控制（MPC）

优势：规划连续无跳变
劣势：优化速度慢，可能无解

3.2.2 规划范式

PLANNING Paradigm

运动规划（motion planning）范例 [28]

Cost functional

$$\text{minimize} \quad J(\boldsymbol{x}(t)) = \int_{t_0}^{t_0+T} L(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}, \dddot{\boldsymbol{x}})\, dt$$

Constraints

$$\text{subject to} \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \quad \text{and} \quad \boldsymbol{x}(t_0+T) \in X_{\text{goal}}$$

$$f(\boldsymbol{x}, \dot{\boldsymbol{x}}, \dots) = 0 \quad \forall t \in [t_0, t_0+T]$$

$$g(\boldsymbol{x}, \dot{\boldsymbol{x}}, \dots) \leq 0 \quad \forall t \in [t_0, t_0+T]$$

Internal $\begin{cases} \text{Kinematic} \\ \text{Dynamic} \end{cases}$

External $\begin{cases} \text{Driving corridor} \\ \text{Obstacles} \end{cases}$

Example:

$$L = \underbrace{w_{\text{off}} \left| \boldsymbol{x} - \boldsymbol{x}_{\text{ref}} \right|^2}_{\text{Follow the reference path}} + \underbrace{w_{\text{vel}} \left| \dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_{\text{ref}} \right|^2}_{\text{Reach the desired velocity}} + \underbrace{w_{\text{acc}} \left| \ddot{\boldsymbol{x}} \right|^2}_{\text{Reduce accelerations}} + \underbrace{w_{\text{jerk}} \left| \dddot{\boldsymbol{x}} \right|^2}_{\text{Reduce jerk}}$$

3.2.3 代价函数

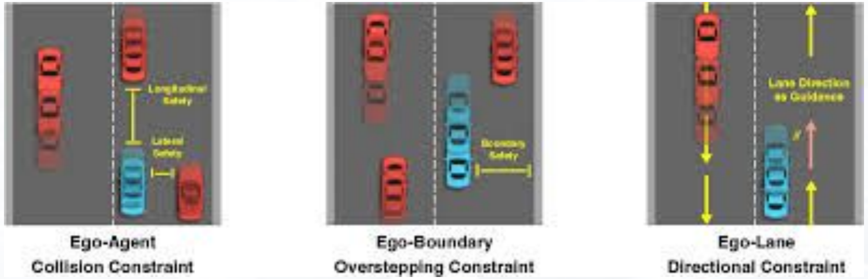
Cost functions

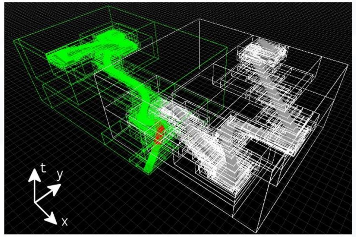
代价函数:对自车规划状态性能的评价 [58]

- 全面性:覆盖更全面的驾驶场景
- 可优化:选择易优化、可微分的函数

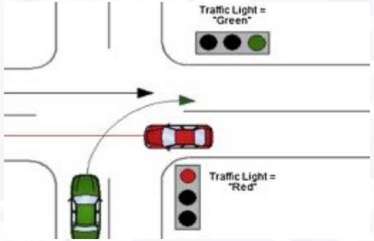
代价函数考虑因素 [38]

- 安全性:同静态,动态实例无碰撞
- 合规性:道路行驶,及交通规则合规
- 舒适性:规划轨迹平滑,驾驶停顿小
- 效率性:驾驶向目标,有持续的规划进度

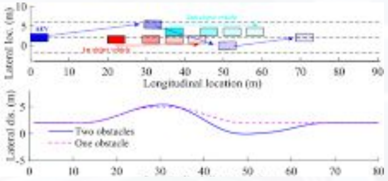




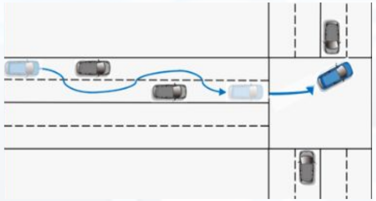
安全性



合规性



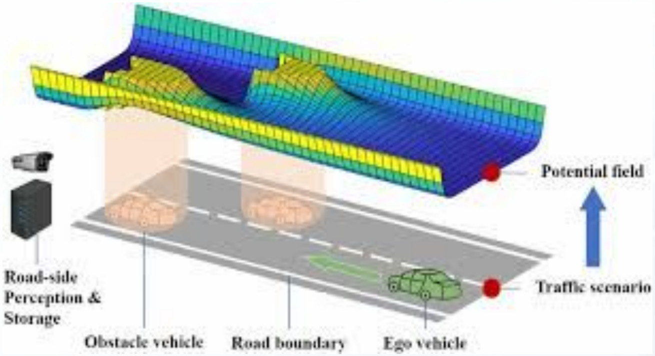
舒适性

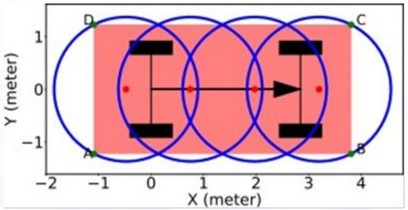


效率性

3.2.3 代价函数
Cost functions

通用代价函数实例





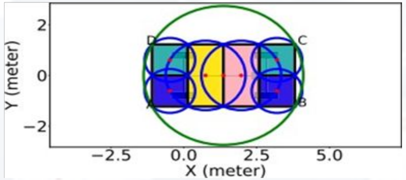
四圆分解(4C decompose)[64]

安全性

- 距离启发 [64]
- 互斥场(repulsive field)[65]

$$\frac{1}{d\left(\hat{\mathbf{s}}_{m,t}^{(i,k)}, \hat{\mathbf{s}}_{n,t}^{(j,k-1)}\right)+1},$$

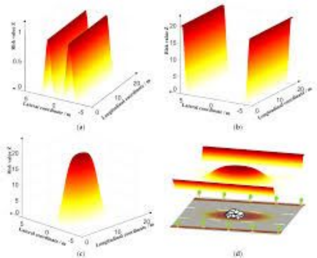
$$U_{OB} = A_{obs}\exp[-\frac{C_1}{2}(\frac{(X-X_o)^2}{\sigma_x}+\frac{(Y-Y_o)^2}{\sigma_y}-C_2)],$$



六圆分解(6C decompose)[64]

合规性

- 中心线(centerline)距离
- 可行域(drivable area)分析 [66]



$$\hat{O}^t_{\text{non-drivable}} = \hat{O}^t_{\text{agent}} | \bar{O}_{\text{static}} | \overline{\bar{O}}_{\text{drivable}},$$

$$\hat{O}^t = \text{Conv}(W(\hat{\tau}_t, H_{ego}, W_{ego}), \bar{O}_{\text{non-drivable}})$$

3.2.3 代价函数

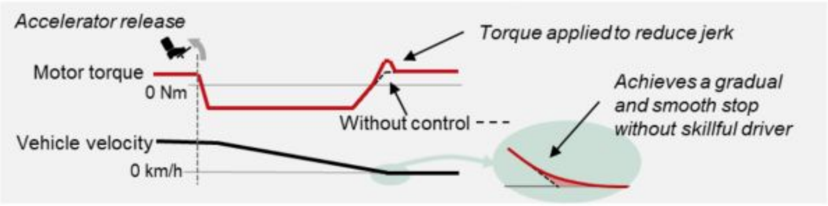
Cost functions

通用代价函数实例

舒适性

- 最小化jerk
- 最小横向加速度

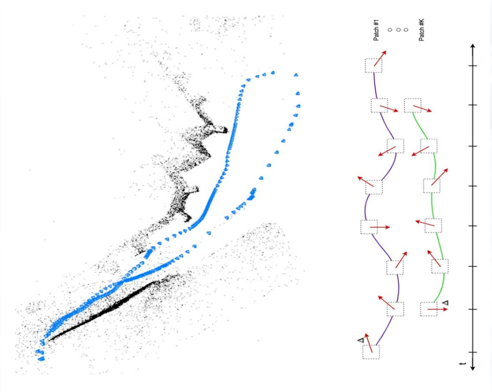
$$w_{\text{jerk}} \left| \dddot{x} \right|^2$$



效率性

- 累计规划距离

$$-\lambda \sum_{t=0}^{T_f} \|\mathbf{w}_t\|_1$$



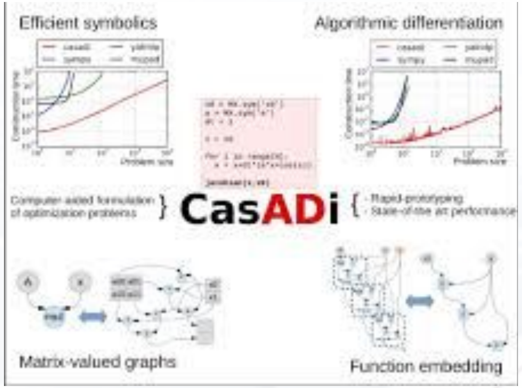
3.2.4 优化器
Optimization Tools

CasADi [67]

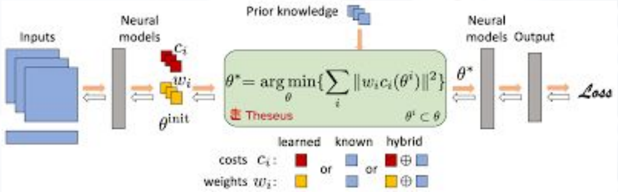
- 领域最通用的优化工具
- 兼容多种程序语言
- 直观的优化流程 (pipeline)

Theseus [68]

- 可微分 (differentiable) 的优化工具
- 兼容Pytorch
- 可嵌入模型及单卡训练环节



CasADi



Theseus

3.3 规划指标

Planning Metrics

规划进度 / 相似度指标

- L2距离
- 道路完成度（Route completion，RC） $\frac{1}{N}\sum_{i}^{N} R_i$
- 专家进度（Expert‘s progress，EP） $\frac{1}{N}\sum_{N}\frac{R_i^{\text{Plan}}}{R_i^{\text{Expert}}}$

安全/合规指标

- 碰撞率 (Collision rate，CR)
- 可行域合规率（Area compliance，AC）
- 违规率（Infraction score，IS）

$$\prod_{j}^{ped.,\ldots,stop}(p_i^j)^{\#infractions}$$

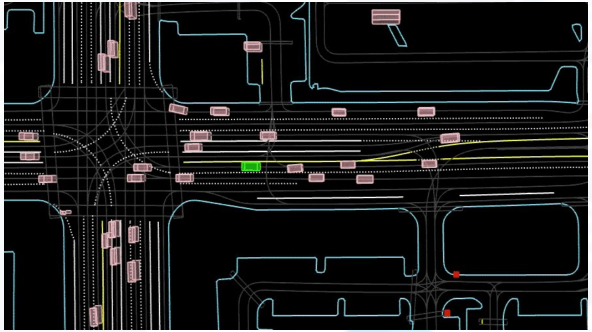
3.4 规划基准
Planning Benchmarks

规划评测范式 [69]

- 开环评测（Open-loop）（不推荐！[70]）
- 闭环回放评测（Closed-loop log-replay）
- 闭环模拟评测（Closed-loop reactive）

规划评测数据集/模拟器

- nuPlan/nuPlan模拟器（>1000,000场景）[69]
- Waymo/Waymax 模拟器 [52, 71]
- CARLA模拟器 [72]



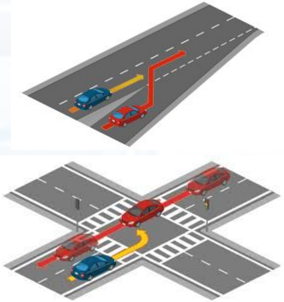
3.4 规划基准
Planning Benchmarks

规划基准:nuPlan

- Test14 [73]: 共约8400场景，3种评测方式
- Val14 [74]: 共约1040场景，3种评测方式
- 排行榜:2023(已停用);2024(端到端)

规划基准:CARLA

- Longest6 [75]: 约36个town，闭环模拟
- Town05 [75]: 一个town，闭环模拟
- 排行榜: (CARLA1，CARLA2)闭环模拟



4.课后思考/Open Questions

4. 课后思考

Open Questions

- 课上学习到不同的预测与规划方法，分别有什么优势和劣势？
- 预测与规划如何进行模块协同和融合（integration），都可能遇到什么问题？
- 预测和规划真的误差越小越好嘛？是否存在误差允许的下界？
- 预测和规划如何选择建模，同上下游感知、控制的任务模块适配？
- 预测任务和世界模型（world model）有什么联系与区别？
- 如何进一步提升验证指标，评测流程和基准（benchmark）的真实性与代表性？

Reference

Part 1

[1] Chen, L., Li, Y., Huang, C., Li, B., Xing, Y., Tian, D., ... & Wang, F. Y. (2022). Milestones in autonomous driving and intelligent vehicles: Survey of surveys. IEEE Transactions on Intelligent Vehicles, 8(2), 1046-1056.

[2] Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., ... & De Souza, A. F. (2021). Self-driving cars: A survey. Expert systems with applications, 165, 113816.

[3] Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabbat, M., ... & Ballas, N. (2023). Self-supervised learning from images with a joint-embedding predictive architecture. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15619-15629).

[4] Chen, C., Seff, A., Kornhauser, A., & Xiao, J. (2015). Deepdriving: Learning affordance for direct perception in autonomous driving. In Proceedings of the IEEE international conference on computer vision (pp. 2722-2730).

[5] Hagedorn, S., Hallgarten, M., Stoll, M., & Condurache, A. (2023). Rethinking integration of prediction and planning in deep learning-based automated driving systems: a review. arXiv preprint arXiv:2308.05731.

[6] Zhou, Z., Wang, J., Li, Y. H., & Huang, Y. K. (2023). Query-centric trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 17863-17873).

[7] Wang, W., Wang, L., Zhang, C., Liu, C., & Sun, L. (2022). Social interactions for autonomous driving: A review and perspectives. Foundations and Trends® in Robotics, 10(3-4), 198-376.

[8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

[9] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. ICLR, 2017.

Reference

Part 2

[10] Mozaffari, S., Al-Jarrah, O. Y., Dianati, M., Jennings, P., & Mouzakitis, A. (2020). Deep learning-based vehicle behavior prediction for autonomous driving applications: A review. IEEE Transactions on Intelligent Transportation Systems, 23(1), 33-47.

[11] Fang, J., Wang, F., Xue, J., & Chua, T. S. (2024). Behavioral intention prediction in driving scenes: A survey. IEEE Transactions on Intelligent Transportation Systems.

[12] Li, L., Zhao, W., Xu, C., Wang, C., Chen, Q., & Dai, S. (2021). Lane-change intention inference based on RNN for autonomous driving on highways. IEEE Transactions on Vehicular Technology, 70(6), 5499-5510.

[13] Fang, J., Yan, D., Qiao, J., Xue, J., & Yu, H. (2021). DADA: Driver attention prediction in driving accident scenarios. IEEE transactions on intelligent transportation systems, 23(6), 4959-4971.

[14] Chen, Y., Veer, S., Karkus, P., & Pavone, M. (2023). Interactive joint planning for autonomous vehicles. IEEE Robotics and Automation Letters.

[15] Phan-Minh, T., Grigore, E. C., Boulton, F. A., Beijbom, O., & Wolff, E. M. (2020). Covernet: Multimodal behavior prediction using trajectory sets. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 14074-14083).

[16] Kim, H., Lee, K., Hwang, G., & Suh, C. (2019, July). Crash to not crash: Learn to identify dangerous vehicles using a simulator. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 978-985).

[17] Banijamali, E., Rohani, M., Amirloo, E., Luo, J., & Poupart, P. (2021). Prediction by anticipation: An action-conditional prediction method based on interaction learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 15621-15630).

[18] Luo, Y., Cai, P., Bera, A., Hsu, D., Lee, W. S., & Manocha, D. (2018). Porca: Modeling and planning for autonomous driving among many pedestrians. IEEE Robotics and Automation Letters, 3(4), 3418-3425.

Reference

Part 2

[19] Berndt, Holger, Jorg Emmert, and Klaus Dietmayer. "Continuous driver intention recognition with hidden markov models." 2008 11th International IEEE Conference on Intelligent Transportation Systems. IEEE, 2008.

[20] Zhang, Y., Lin, Q., Wang, J., Verwer, S., & Dolan, J. M. (2018). Lane-change intention estimation for car-following control in autonomous driving. IEEE Transactions on Intelligent Vehicles, 3(3), 276-286.

[21] Liu, J., Luo, Y., Zhong, Z., Li, K., Huang, H., & Xiong, H. (2022). A probabilistic architecture of long-term vehicle trajectory prediction for autonomous driving. Engineering, 19, 228-239.

[22] Yau, T., Malekmohammadi, S., Rasouli, A., Lakner, P., Rohani, M., & Luo, J. (2021, May). Graph-sim: A graph-based spatiotemporal interaction modelling for pedestrian action prediction. In 2021 IEEE International Conference on Robotics and Automation (ICRA) (pp. 8580-8586). IEEE.

[23] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).

[24] Huang, Y., Du, J., Yang, Z., Zhou, Z., Zhang, L., & Chen, H. (2022). A survey on trajectory-prediction methods for autonomous driving. IEEE Transactions on Intelligent Vehicles, 7(3), 652-674.

[25] Wiest, J., Höffken, M., Kreßel, U., & Dietmayer, K. (2012, June). Probabilistic trajectory prediction with Gaussian mixture models. In 2012 IEEE Intelligent vehicles symposium (pp. 141-146). IEEE..

[26] Varadarajan, B., Hefny, A., Srivastava, A., Refaat, K. S., Nayakanti, N., Cornman, A., ... & Sapp, B. (2022, May). Multipath++: Efficient information fusion and trajectory aggregation for behavior prediction. In 2022 International Conference on Robotics and Automation (ICRA) (pp. 7814-7821). IEEE.

[27] Zhang, L., Li, P., Liu, S., & Shen, S. (2024). SIMPL: A Simple and Efficient Multi-agent Motion Prediction Baseline for Autonomous Driving. IEEE Robotics and Automation Letters.

Reference

Part 2

[28] Lecture Notes “Vehicle Dynamics of Passenger Cars”, Institute of Automotive Technology, TUM

[29] R. Schubert, E. Richter, and G. Wanielik, “Comparison and evaluation of advanced motion models for vehicle tracking,” 2008 11th International Conference on Information Fusion, pp. 1–6, 2008.

[30] Jia, X., Chen, L., Wu, P., Zeng, J., Yan, J., Li, H., & Qiao, Y. (2023, March). Towards capturing the temporal dynamics for trajectory prediction: a coarse-to-fine approach. In Conference on Robot Learning (pp. 910-920). PMLR.

[31] Zhou, Z., Ye, L., Wang, J., Wu, K., & Lu, K. (2022). Hivt: Hierarchical vector transformer for multi-agent motion prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8823-8833).

[32] Chai, Y., Sapp, B., Bansal, M., & Anguelov, D. (2019). Multipath: Multiple probabilistic anchor trajectory hypotheses for behavior prediction. arXiv preprint arXiv:1910.05449.

[33] Huang, R., Xue, H., Pagnucco, M., Salim, F., & Song, Y. (2023). Multimodal trajectory prediction: A survey. arXiv preprint arXiv:2302.10463.

[34] Gao, J., Sun, C., Zhao, H., Shen, Y., Anguelov, D., Li, C., & Schmid, C. (2020). Vectornet: Encoding hd maps and agent dynamics from vectorized representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 11525-11533).

[35] Mo, X., Huang, Z., Xing, Y., & Lv, C. (2022). Multi-agent trajectory prediction with heterogeneous edge-enhanced graph attention network. IEEE Transactions on Intelligent Transportation Systems, 23(7), 9554-9567.

[36] Sun, Q., Huang, X., Gu, J., Williams, B. C., & Zhao, H. (2022). M2i: From factored marginal trajectory prediction to interactive prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6543-6552).

Reference

Part 2

[37] Huang, Z., Mo, X., & Lv, C. (2022, May). Multi-modal motion prediction with transformer-based neural network for autonomous driving. In 2022 International Conference on Robotics and Automation (ICRA) (pp. 2605-2611). IEEE.

[38] Zhao, H., Gao, J., Lan, T., Sun, C., Sapp, B., Varadarajan, B., ... & Anguelov, D. (2021, October). Tnt: Target-driven trajectory prediction. In Conference on Robot Learning (pp. 895-904). PMLR.

[39] Zhang, L., Su, P. H., Hoang, J., Haynes, G. C., & Marchetti-Bowick, M. (2021, October). Map-adaptive goal-based trajectory prediction. In Conference on Robot Learning (pp. 1371-1383). PMLR.

[40] Gilles, T., Sabatini, S., Tsishkou, D., Stanciulescu, B., & Moutarde, F. (2021, September). Home: Heatmap output for future motion estimation. In 2021 IEEE International Intelligent Transportation Systems Conference (ITSC) (pp. 500-507). IEEE.

[41] Phan-Minh, T., Grigore, E. C., Boulton, F. A., Beijbom, O., & Wolff, E. M. (2020). Covernet: Multimodal behavior prediction using trajectory sets. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 14074-14083).

[42] Cui, H., Radosavljevic, V., Chou, F. C., Lin, T. H., Nguyen, T., Huang, T. K., ... & Djuric, N. (2019, May). Multimodal trajectory predictions for autonomous driving using deep convolutional networks. In 2019 international conference on robotics and automation (ICRA) (pp. 2090-2096). IEEE.

[43] Hu, A., Murez, Z., Mohan, N., Dudas, S., Hawke, J., Badrinarayanan, V., ... & Kendall, A. (2021). Fiery: Future instance prediction in bird's-eye view from surround monocular cameras. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 15273-15282).

[44] Agro, B., Sykora, Q., Casas, S., & Urtasun, R. (2023). Implicit occupancy flow fields for perception and prediction in self-driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1379-1388).

Reference

Part 2

[45] Hoermann, S., Bach, M., & Dietmayer, K. (2018, May). Dynamic occupancy grid prediction for urban autonomous driving: A deep learning approach with fully automatic labeling. In 2018 IEEE International Conference on Robotics and Automation (ICRA) (pp. 2056-2063). IEEE.

[46] U.S. Department of Transportation Federal Highway Administration. (2016). Next Generation Simulation (NGSIM) Vehicle Trajectories and Supporting Data.

[47] Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018, November). The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In 2018 ITSC (pp. 2118-2125). IEEE.

[48] Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., ... & Beijbom, O. (2020). nuscenes: A multimodal dataset for autonomous driving. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11621-11631).

[49] Zhan, W., Sun, L., Wang, D., Shi, H., Clausse, A., Naumann, M., ... & Tomizuka, M. (2019). Interaction dataset: An international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps. arXiv preprint arXiv:1910.03088.

[50] Chang, M. F., Lambert, J., Sangkloy, P., Singh, J., Bak, S., Hartnett, A., ... & Hays, J. (2019). Argoverse: 3d tracking and forecasting with rich maps. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8748-8757).

[51] Wilson, B., Qi, W., Agarwal, T., Lambert, J., Singh, J., Khandelwal, S., ... & Hays, J. (2023). Argoverse 2: Next generation datasets for self-driving perception and forecasting. arXiv preprint arXiv:2301.00493.

[52] Ettinger, S., Cheng, S., Caine, B., Liu, C., Zhao, H., Pradhan, S., ... & Anguelov, D. (2021). Large scale interactive motion forecasting for autonomous driving: The waymo open motion dataset. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 9710-9719).

Reference

Part 3

[53] Teng, S., Hu, X., Deng, P., Li, B., Li, Y., Ai, Y., ... & Chen, L. (2023). Motion planning for autonomous driving: The state of the art and future perspectives. IEEE Transactions on Intelligent Vehicles.

[54] Xu, W., Wei, J., Dolan, J. M., Zhao, H., & Zha, H. (2012, May). A real-time motion planner with trajectory optimization for autonomous vehicles. In 2012 IEEE International Conference on Robotics and Automation (pp. 2061-2067). IEEE.

[55] Werling, M., Ziegler, J., Kammel, S., & Thrun, S. (2010, May). Optimal trajectory generation for dynamic street scenarios in a frenet frame. In 2010 IEEE international conference on robotics and automation (pp. 987-993). IEEE.

[56] Chen, J., Yuan, B., & Tomizuka, M. (2019, November). Deep imitation learning for autonomous driving in generic urban scenarios with enhanced safety. In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 2884-2890). IEEE.

[57] Cheng, J., Chen, Y., & Chen, Q. (2024). PLUTO: Pushing the Limit of Imitation Learning-based Planning for Autonomous Driving. arXiv preprint arXiv:2404.14327.

[58] Claussmann, L., Revilloud, M., Gruyer, D., & Glaser, S. (2019). A review of motion planning for highway autonomous driving. IEEE Transactions on Intelligent Transportation Systems, 21(5), 1826-1848.

[59] Hubmann, C., Schulz, J., Becker, M., Althoff, D., & Stiller, C. (2018). Automated driving in uncertain environments: Planning with interaction and uncertain maneuver prediction. IEEE transactions on intelligent vehicles, 3(1), 5-17.

[60] Huang, Z., Wu, J., & Lv, C. (2021). Driving behavior modeling using naturalistic human driving data with inverse reinforcement learning. IEEE transactions on intelligent transportation systems, 23(8), 10239-10251.

Reference

Part 3

[61] Chen, Y., Karkus, P., Ivanovic, B., Weng, X., & Pavone, M. (2023, May). Tree-structured policy planning with learned behavior models. In 2023 IEEE International Conference on Robotics and Automation (ICRA) (pp. 7902-7908). IEEE.

[62] Chu, K., Lee, M., & Sunwoo, M. (2012). Local path planning for off-road autonomous driving with avoidance of static obstacles. IEEE transactions on intelligent transportation systems, 13(4), 1599-1616.

[63] Liu, C., Lee, S., Varnhagen, S., & Tseng, H. E. (2017, June). Path planning for autonomous vehicles using model predictive control. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 174-179). IEEE.

[64] Ericson, C. (2004). Real-time collision detection. Crc Press.

[65] Hwang, Y. K., & Ahuja, N. (1992). A potential field approach to path planning. IEEE transactions on robotics and automation, 8(1), 23-32.

[66] Hu, Y., Li, K., Liang, P., Qian, J., Yang, Z., Zhang, H., ... & Liu, Q. (2023). Imitation with spatial-temporal heatmap: 2nd place solution for nuplan challenge. arXiv preprint arXiv:2306.15700.

[67] Andersson, J. A., Gillis, J., Horn, G., Rawlings, J. B., & Diehl, M. (2019). CasADi: a software framework for nonlinear optimization and optimal control. Mathematical Programming Computation, 11, 1-36.

[68] Pineda, L., Fan, T., Monge, M., Venkataraman, S., Sodhi, P., Chen, R. T., ... & Mukadam, M. (2022). Theseus: A library for differentiable nonlinear optimization. Advances in Neural Information Processing Systems, 35, 3801-3818.

Reference

Part 3

[69] Caesar, H., Kabzan, J., Tan, K. S., Fong, W. K., Wolff, E., Lang, A., ... & Omari, S. (2021). nuplan: A closed-loop ml-based planning benchmark for autonomous vehicles. arXiv preprint arXiv:2106.11810.

[70] Li, Z., Yu, Z., Lan, S., Li, J., Kautz, J., Lu, T., & Alvarez, J. M. (2024). Is Ego Status All You Need for Open-Loop End-to-End Autonomous Driving?. CVPR, 2024.

[71] Gulino, C., Fu, J., Luo, W., Tucker, G., Bronstein, E., Lu, Y., ... & Sapp, B. (2024). Waymax: An accelerated, data-driven simulator for large-scale autonomous driving research. Advances in Neural Information Processing Systems, 36.

[72] Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., & Koltun, V. (2017, October). CARLA: An open urban driving simulator. In Conference on robot learning (pp. 1-16). PMLR.

[73] Cheng, J., Chen, Y., Mei, X., Yang, B., Li, B., & Liu, M. (2024). Rethinking Imitation-based Planner for Autonomous Driving. ICRA, 2024.

[74] Dauner, D., Hallgarten, M., Geiger, A., & Chitta, K. (2023, December). Parting with misconceptions about learning-based vehicle motion planning. In Conference on Robot Learning (pp. 1268-1281). PMLR.

[75] Prakash, A., Chitta, K., & Geiger, A. (2021). Multi-modal fusion transformer for end-to-end autonomous driving. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 7077-7087).

End