# 决策算法

## 决策问题

### 典型场景

考虑如下两个典型场景：

1. 在正常道路上，自车(代号为0)一开始正在中间道路上行驶，打算向右侧变道，那么应该是加速在5号车前面变道过去还是减速在后方变道过去？又或者打算向左变道，那么应该何时加塞进入到左侧车道呢？



1. 自车准备左转无保护左转时，应该如何动作才能顺利进入左转车道呢？



从上面两个场景来看，自车需要对当前场景进行分析，预测他车的可能的行为，并根据他车的行为做出自车应该的行为。所以这里问题分解为：1. 通过自车和他车的历史轨迹信息估计未来可能的行为。2. 利用估计的行为给出自车的大致行为。

## 2. 决策过程

### 2.1 横纵向行为分析

我们将横向行为表示为左转、居中、右转。纵向行为分为加速、减速、保持。假设车辆行为的变化周期为,我们整个规划周期为,那么在整个规划周期内，可以规划次，假设n = 2,来看一下会存在多少种可能性？



可以从上图分析看出，即使n=2，那么可能的操作空间为9\*9 = 81次，那么n次，便有条分支，如果场景中有m辆车，那么个分支，显然这样的分支过于庞大，cpu很难在短时间内将所有分支遍历完成，因此这里需要对分支进行裁剪。

### 裁剪

#### 他车分支裁剪

我们假设，其他agents在整个决策周期内，其可能的意图不会发生变化，也就是说，在一开始我们通过他车的历史轨迹信息推测出其他agents可能的意图在整个决策周期内不会发生变化。这样做实际上问题不大，原因是，首先我们决策每0.1s执行一次，也就是说每0.1s我们会对他车的行为重新进行预测，从响应频率上是满足的。这样做，在损失较小代价下换来了时间上的效率。这样，原来的分支便缩小为条，即使这样，分支还是过多，因此需要对自车进行裁剪。

#### 自车分支裁剪

对于自车来说，我们假设自车在整个规划周期内横向上最多只做一次变化，纵向上最多变化一次，不过横向和纵向上还是有些许区别，具体看下图。







在纵向上，我们设计为从一开始就确定纵向行为，在横向上则是可能在不同时刻发生行为变化，一旦横向行为变化，后续便不在变化。这样行为分支大致为条。通过以上的剪枝策略后，接下来就是他车的行为预测和具体的仿真模拟了。

### 他车行为预测

通常对于他车的预测，有一个专门的模块。但是在这里，我们将预测分为两步，第一步通过历史信息分析出接下来他车可能的横向和纵向行为。第二步，通过仿真模拟，与自车一起实时的计算出具体的轨迹信息。这样做的一个好处是，因为不管是自车还是他车的意图都是需要实时交互的，或者说需要博弈的，而不是一开始就是确定的。这里首先分析的是他车可能的横向或纵向行为，这里有两种方法，第一中基于规则的方法，第二种基于学习的方法，第一种通常较为简单，通过简单的规则逻辑确定行为，容易快速实现，但是正确性上面可能会差一些，或者说需要更复杂的判断才能提高准确性，第二中，需要数据驱动，工程上更复杂，需要大量数据做支撑。这里重点分析具有规则的，简单介绍一些介于学习的。

#### 基于规则的行为预测

对于任意一辆车来说，我们判断它可能的横向行为的概率。这里我们引用文献【1】的预测方式进行行为预测。

**The lane-changing model MOBIL**

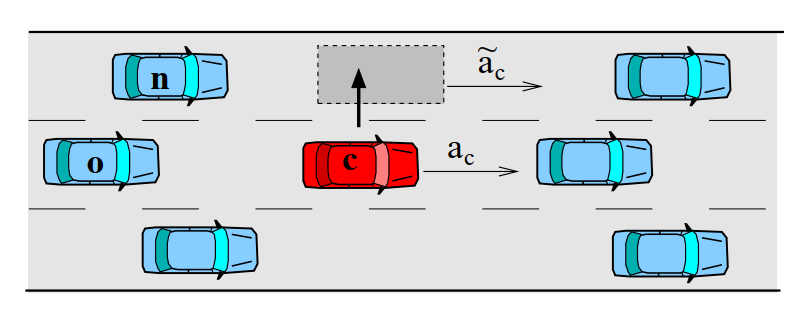


图1

如图所示，考虑自车c准备向左侧进行变道，车o和车n分别是当前车道和目标车道的follow car.那么car-following 模型表示为：

其中，表示自车与前车的距离，表示自车的速度，

**Safety Criterion**

为了变道的安全性，引入安全准则。通过考虑对目标车道上上游车辆的影响来检查执行变道(间隙接受)的可能性。安全准则保证，在变道后，目标车道上后继的加速度 不应该超过给定的安全速度限制

尽管这个公式非常简单，但是这个条件已经包括了纵向car-following模型的所有信息，应为本身已经依赖了自车速度、前车距离等等。当目标车道上的跟车速度比自己的速度快时，需要后车与自己的位置有较大的距离才能满足安全约束。如果跟车速度较慢，那么间隙可以小一点。对于真实的纵向模型，应该低于最大可能的减速度.

**Incentive criterion for symmetric(‘U’S) lane-changes rules**

激励标准通常决定变道是否改善了驾驶员的个人当地交通状况。在我们的模型中，我们将激励标准推广到包括直接受影响的相邻车辆。礼貌因子p决定了这些车辆对变道决策的影响程度。于是有：

表示变道对自车的效用，则表示对当前后车和变道后的后车的效用。

右侧的模拟了一定的惯性，防止频繁换道。综上所述，如果自身的优势（加速增益）大于新旧后继者的劣势（加速损失）与阈值的加权和，则满足激励标准。阈值影响全局变道行为，而礼貌因子影响局部行为即新旧后继者。

由于其他司机的劣势和自己的优势通过礼貌因素p得到平衡，因此变道模型具有典型的经典博弈策略特征。P值可以解释为利他主义的程度。它可以从 p = 0（对于自私的换道者）到 p > 1（对于无私的司机），对于无私的司机，如果换道导致followers的交通情况变糟糕，则不会考虑换道，反而如果可以提高followers的效率，即使存在不利因素也会变道。通过设置 p < 0，甚至可以对恶意驾驶员进行建模，他们接受自己的劣势以阻止他人。

特殊情况下，p=1且,激励准则简化为：

因此，只有当所有相关车辆的加速度总和增加时，才会执行变道，这符合理想意义上的“最小化变道引起的整体制动”（MOBIL）的概念。在这种情况下，不需要额外的安全约束，因为只要加速度方面的优势低于制动减速度方面的劣势，方程（4）就会自动排除为了避免事故而进行的制动操作。因此，对应于 p = 1 的“理想”MOBIL 策略没有自由参数，因此可能被视为换道决策的“最小模型”。 在第 3 节中，我们研究了车道变更率（每公里和每小时），该比率主要由礼貌因素 p 决定。

由此我们可以给出变道行为的预测流程图。

****

图2

#### 基于学习的行为预测

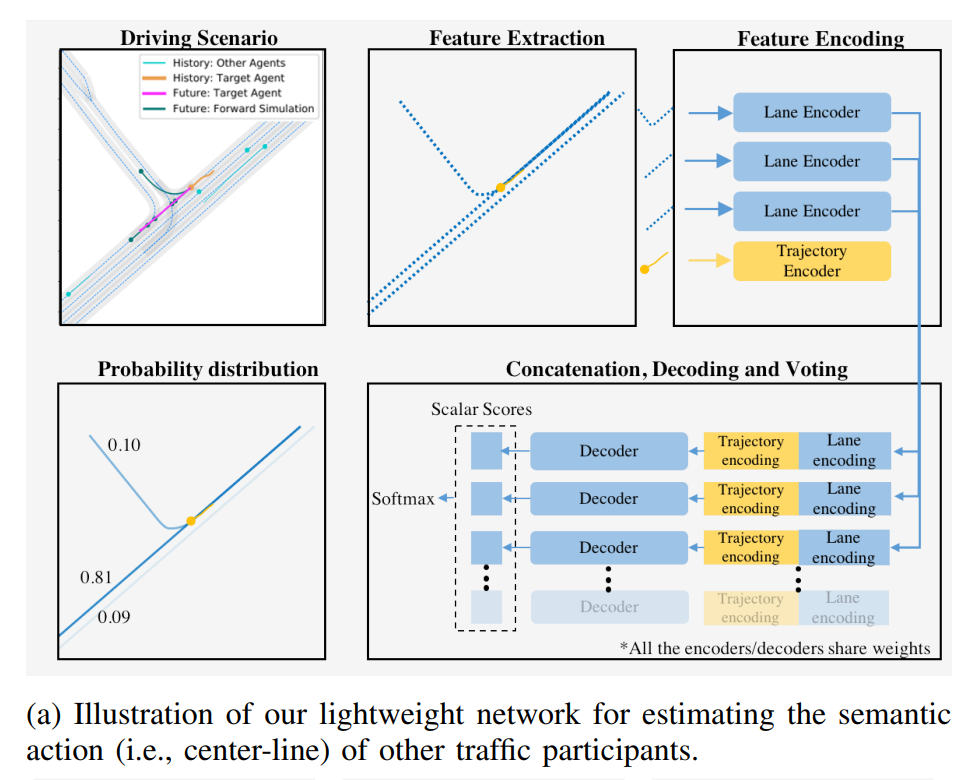


图3

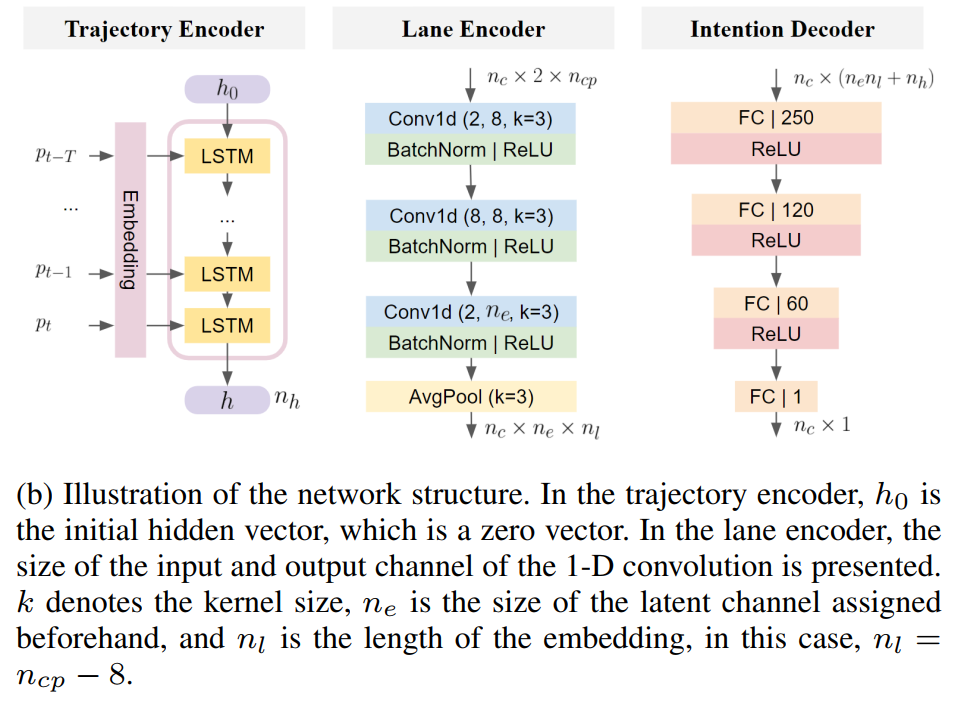


图4

网络结构如图所示，在顶层，对候选的中心线以及自车的历史轨迹进行采样，作为输入feature并且使用encoder-decoder结构对每条中心线进行打分预测，并最后由softmax进

行正则化输出。由于该网络预测每个中心线实例的标量分数，并且不依赖于场景中中心线nc的数量，因此该网络可以在各种城市环境中工作，如图5所示。

如图4所示，该网络有三个主要模块，lane encoder:将周围环境的所有车道中线进行编码，对于每个中心线采样70个点，中心线特征首先用核大小为3的多个一维卷积层进行编码，再用average pooling。对于目标车辆的轨迹，我们使用LSTM对二维位置进行循环编码，并产生一个潜在向量，它是运动历史的紧凑表示。设置LSTM的输入尺寸为4，隐藏层的深度为64.对车道和车辆轨迹进行处理后，我们将轨迹嵌入与每个车道嵌入连接起来，然后使用带有ReLU激活函数的多层感知器(MLP)来估计候选车道的分数，最后的概率分布通过softmax函数得到。由于意图预测本质上是一个分类问题，我们使用负对数似然(NLL)作为损失最小化。

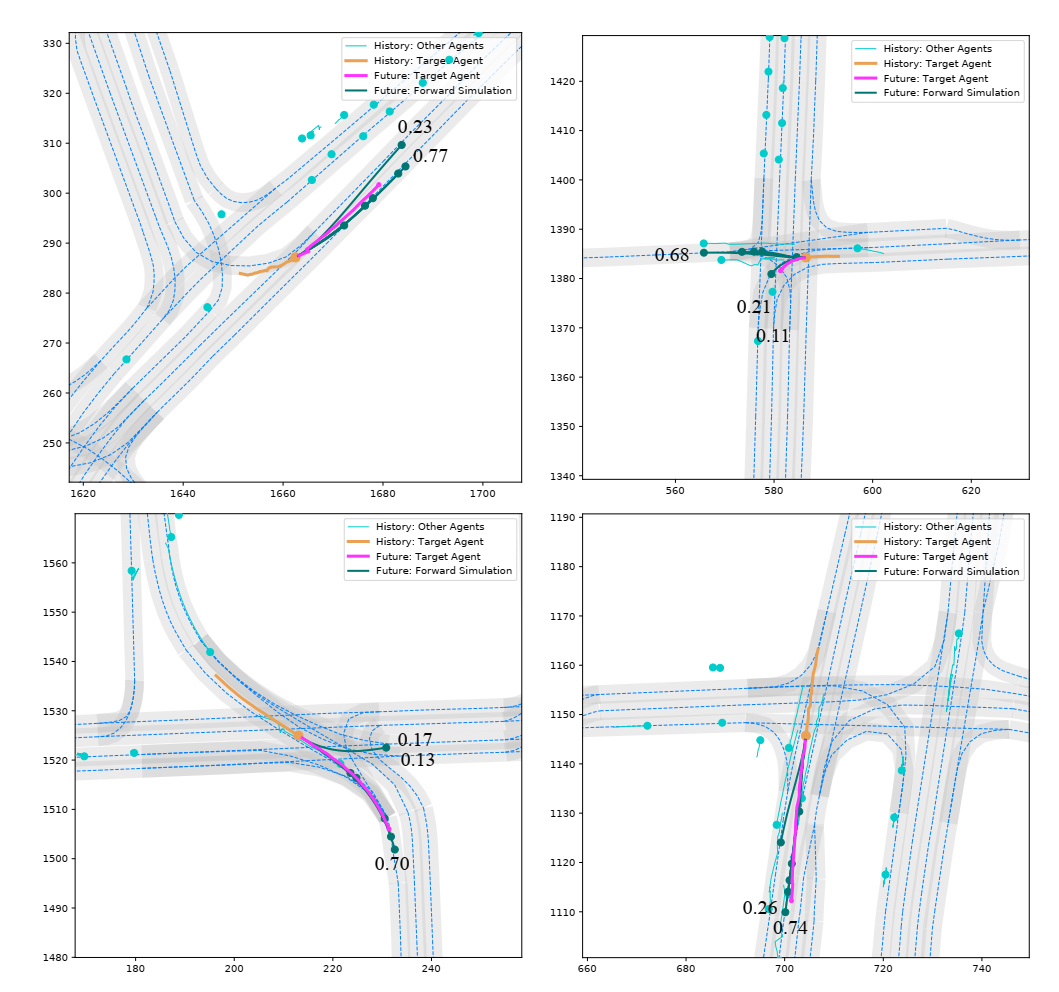


图5

### Simulator

#### 2.4.1 IDM模型

## 参考文献

[1] Kesting A .General lane-changing model MOBIL for car-following models; ; Transportation research record; Traffic flow theory 2006[J]. 2007.