

DOI: 10.19650/j.cnki.cjsi.J2413641

基于大语言模型的可解释性运动规划方法研究*

陈熙源, 刘炜焱, 聂姝涵, 经纬铭

(东南大学仪器科学与工程学院 南京 210096)

摘要:基于学习的运动规划方法采用数据驱动的策略,从大规模的驾驶经验中学习策略,虽能表现出良好的性能,但因将运动规划视为黑箱问题,牺牲了方法的可解释性,也总会遇到数据集偏差、过拟合以及陷入局部最优等挑战。利用新兴的大语言模型强大的推理能力和解释能力,提出了一个基于大语言模型的自动驾驶运动规划框架,称为 LLMs-Driver,来解决基于学习的方法中可解释性差的问题。LLMs-Driver 由推理模块、记忆模块和反思模块 3 部分组成。在推理模块中,提出了重要经验回放算法,综合考虑了经验优先级和场景相似性两个影响因素,以提高 LLMs-Driver 的学习效率和性能。在记忆模块中,提出了改进后的先进先出经验存储算法,以保证经验的有效性和新颖性,有助于 LLMs-Driver 学习到最新最好的策略。同时,为了充分增强自动驾驶运动规划模型的透明度和可信度,采用“三步思维链”的方法,将推理和反思过程分别划分为 3 个步骤并配有解释性文字。最后,在 Highway-env 仿真平台上对 LLMs-Driver 进行闭环自动驾驶实验。实验结果表明,LLMs-Driver 有着显著的可解释性和运动规划能力,任务成功步数的中值最高提升至基线算法的 2.19 倍,并支持根据驾驶人的意图设置不同的驾驶风格。

关键词: 自动驾驶; 运动规划; 大语言模型; 重要经验回放算法

中图分类号: TH89 **文献标识码:** A **国家标准学科分类代码:** 520.20 510.80

Research on interpretable motion planning methods with large language models

Chen Xiyuan, Liu Weiyan, Nie Shuhan, Jing Weiming

(School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: The learning-based motion planning approach uses a data-driven policy trained on large-scale driving experiences and have demonstrated good performance. However, these methods often treat motion planning as a black-box problem, resulting in limited interpretability. They also face challenges such as dataset bias, overfitting, and getting stuck in a local optimum. In this paper, we exploit the powerful inference and interpretation capabilities of emerging large language models to propose a large language model-based motion planning framework for autonomous driving, called LLMs-Driver, to address the problem of poor interpretability in learning-based approaches. LLMs-Driver consists of three parts, namely, the reasoning module, the memory module, and the reflection module. In the reasoning module, we propose the important experience playback algorithm, which integrates the two influencing factors of experience priority and scene similarity, to improve the learning efficiency and performance of the LLMs-Driver. In the memory module, we propose an improved first-in-first-out experience storage algorithm to ensure the validity and novelty of the experience, ensuring that LLMs-Driver continuously learns from the most recent and effective strategies. Meanwhile, in order to fully enhance the transparency and credibility of the self-driving motion planning model, the ‘three-step chain of thought’ method is adopted, which divides the inference and reflection process into three steps, each accompanied by explanatory textual reasoning. Finally, we validate LLMs-Driver through closed-loop autonomous driving experiments on the Highway-env simulation platform. The experimental results show that LLMs-Driver has significant interpretability and motion planning capabilities, with the median number of successful steps on a task increased up to 2.19 times of the baseline algorithm. Additionally, it supports the customization of different driving styles based on the driver’s intention.

Keywords: autonomous driving; motion planning; large language models; the important experience playback algorithm

0 引言

自动驾驶技术,作为现代科技领域中最具发展前景与挑战性的前沿方向之一,其目标在于实现全球交通系统的革命性变革^[1]。在自动驾驶领域中,运动规划的概念占据了核心地位,构成了自动驾驶技术的理论基石,旨在为自动驾驶车辆设计出既安全又舒适的驾驶轨迹^[2-5]。运动规划的复杂性源于其必须适应不断变化的交通环境,并在此过程中做出合理的驾驶决策^[6]。随着自动驾驶汽车与多样化的环境因素及不可预测的人类驾驶行为进行交互,运动规划的稳健性和可解释性对于确保驾驶过程的安全性和可靠性具有至关重要的影响^[7]。

随着基于学习的算法不断进步,如强化学习、模仿学习等,运动规划技术得到了显著的进步。基于学习的运动规划方法采用数据驱动的策略,从大规模的驾驶经验中学习策略,最终可以表现出良好的性能^[8]。比如说 Liu 等^[9]提出了一种基于长短期记忆网络的深度强化学习算法,以乘客舒适度和燃油低消耗为目标,实现了同时兼顾换道效率和车辆安全的功能。但这类方法将运动规划视为黑箱问题,牺牲了方法的可解释性,也总会导致数据集偏差、过拟合以及陷入局部最优解等问题。因此,深入研究如何解决这些问题,不仅能够加深对驾驶场景的理解,还有助于构建更为合理的决策机制,从而有可能提高自动驾驶系统的性能。

新兴的大语言模型 (large language models, LLMs) 具有强大的概括能力、推理能力和解释能力,在解决自动驾驶领域目前的问题方面极具潜力。已有部分学者尝试将 LLMs 用于自动驾驶的运动规划模块,并取得了一定的成果,比如说 Fu 等^[10]利用 LLMs 进行推理,并将推理转化为可操作的驾驶行为,显示了 LLMs 在增强自动驾驶规划方面的多功能性。Wen 等^[11]提出了的卢 (DiLu) 框架,利用知识驱动在自动驾驶中完成运动规划任务。Cui 等^[12]提出了一个“以人为本”的自动驾驶集成框架,使用 LLMs 描述对环境的感知。然而,上述方法的研究均是在单一的驾驶意图下进行的。但在实际的驾驶场景中,驾驶风格会根据驾驶意图不同而改变,相同场景下不同的驾驶风格也会产生不同的决策结果。同时,虽有少数方法设计了记忆模块以提高 LLMs 推理和决策的准确性,但忽略了每条经验本身的价值,没有充分利用高价值的经验数据。

基于上述研究,本研究提出了一个基于 LLMs 的自动驾驶运动规划框架,称为大语言模型驾驶员 (large language models for driver, LLMs-Driver),旨在进一步探索利用 LLMs 完成运动规划任务最有效的方式,充分发

挥 LLMs 强大的概括能力、推理能力和解释能力。具体来说,LLMs-Driver 对实际驾驶场景的意图进行总结和分类,支持根据驾驶员的意图选择不同的驾驶风格。整个框架由推理模块、记忆模块和反思模块 3 部分组成。在推理模块中,提出了重要经验回放算法,该算法综合考虑了经验优先级和场景相似性两个影响因素,先通过场景相似度计算选取与当前状态相似的场景,再根据选取经验的优先级进行采样,以实现选取综合价值最高的经验输入到 LLMs 中,从而有助于 LLMs-Driver 充分学习样本经验,更好地制定最终的策略,提高运动规划的成功率。在记忆模块中,提出了改进后的先进先出经验存储算法,当记忆模块的存储容量达到上线后,用新形成的经验替换掉长时间未使用到的经验,使记忆模块中的经验可以保持新颖性,有助于 LLMs-Driver 学习到最新最好的策略。同时,为了充分增强自动驾驶运动规划模型的透明度和可信度,细化了 LLMs 的推理和反思过程,采用了“三步思维链”的方法,将推理和反思过程分别划分为 3 个步骤,并为每个步骤中配有解释性文字。最后,在 Highway-env 仿真平台上对 LLMs-Driver 进行闭环自动驾驶训练和测试。实验结果表明,LLMs-Driver 有着显著的可解释性和运动规划能力,同时不同的驾驶风格下 LLMs-Driver 的决策结果也存在差异性。

本研究重点研究了自动驾驶车辆如何利用 LLMs 强大的概括能力、推理能力和解释能力,完成指定驾驶风格下的运动规划任务。

1 大语言模型驾驶员

1.1 总体框架

LLMs-Driver 是一个基于 LLMs 且具备不同驾驶风格的运动规划器,可根据当前驾驶的意图选择对应的驾驶风格。整个运动规划器框架由 3 个模块组成,分别为推理模块、记忆模块和反思模块,如图 1 所示。推理模块首先从环境中获取当前状态的观测信息,并通过场景编码器将数据信息转换为描述性语言。接着使用重要经验回放算法从记忆模块中提取的综合评分靠前的经验描述,并与当前场景描述相结合,集成到提示生成器中并输入到 LLMs,由 LLMs 按照推理模块中的“三步思维链”进行推理,输出当前场景中的响应动作。记忆模块中存储了过往驾驶场景中的经验,包括决策提示、推理过程和其他有价值的信息。经验按照价值的不同被赋予不同的优先级,优先级高的经验将有更大的概率被选取输入到 LLMs 中。反思模块用于评估过去的决策序列,并将过去的决策序列分为安全或不安全两类。不安全的决策将采用反思模块中的“三步思维链”修改决策,并将改进后的决策存入记忆模块中。下面将具体介绍推理模块、记忆模块

和反思模块的实现方式,以及 LLMs-Driver 提供的不同的驾驶风格。

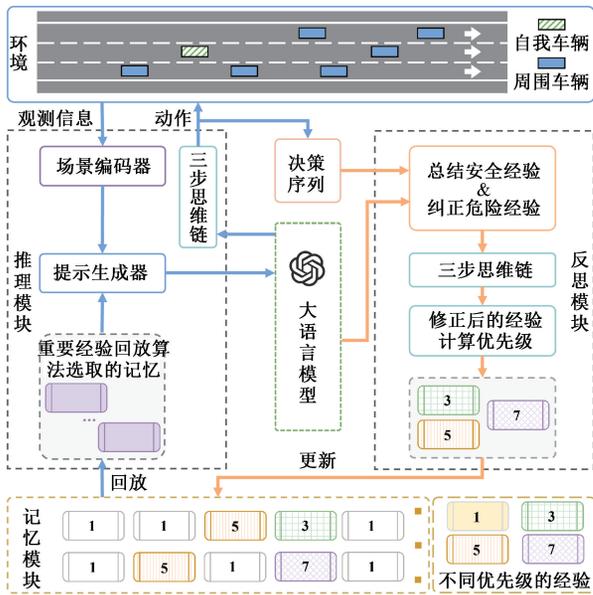


图 1 LLMs-Driver 的整体框架结构

Fig. 1 Overall framework structure of LLMs-Driver

1.2 推理模块

在推理模块中,利用从记忆模块中获得的经验和 LLMs 的常识性知识对当前的交通场景进行决策。具体来说,推理过程如图 2 所示,主要包括:1)场景编码器:通过场景编码器对环境中的观测信息进行编码;2)记忆检索:利用重要经验回放算法从记忆模块中提取综合评分最高的经验;3)生成提示;4)LLMs 在推理模块中的“三步思维链”方法。

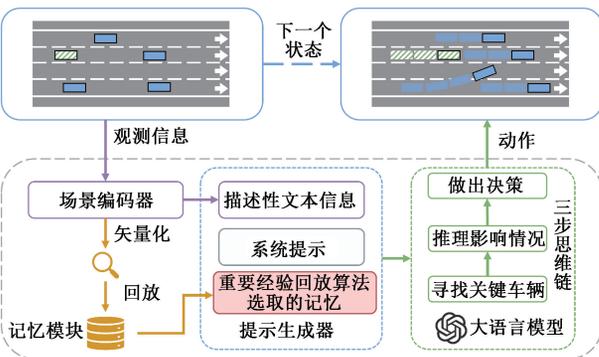


图 2 推理模块结构

Fig. 2 Reasoning module structure

1) 场景编码器

在每个时间步,推理模块会从环境中获取当前时刻自我车辆(ego vehicle, EV)和周围车辆的动态观测信息,

以及道路的静态观测信息。为了便于 LLMs-Driver 了解当前的交通状况,场景描述符将当前场景数据转换为描述性文本。场景描述符遵循标准句子结构,并利用自然语言全面描述正在进行的驾驶场景。例如,交通环境是一个拥有 4 条道路的高速公路,EV 正在驾驶在第 2 个道路上,周围也有其他车辆正在行驶。从环境中可以获取 EV 和周围其他车辆的状态信息,包括位置、速度和加速度。通过描述符对场景进行编码,将场景数据转换为描述性文本,如图 3 所示。最后,这些描述性文本被输入到提示生成器中,作为 LLMs 生成提示语的重要组成部分之一。



图 3 推理模块中的场景编码器

Fig. 3 Scene encoder in the reasoning module

2) 记忆检索

在前人的研究中,DiLu 模型采用相似性查询的方式,选取与当前场景相似的历史经验并输入到提示生成器中。但这种方式忽略了每条经验本身的价值,没有充分利用高价值的经验数据,同时只关注和当前场景相似的经验也容易限于局部最优的困境中。

针对此问题,提出了重要经验回放算法,对从记忆模块中选取经验的方式进行了改进,伪代码如算法 1 所示。重要经验回放算法中综合考虑了经验优先级和场景相似度两个因素。首先,先利用场景相似度的计算公式提取出与当前场景较为相似的若干个经验,再利用优先级对这些经验进行排序,选出对当前决策最有价值的几组经验。在推理的过程中,将当前驾驶场景的描述性文本信息嵌入到一个向量中,再利用场景相似度搜索前 k 个与该场景相似的记忆经验。

算法 1: 重要经验回放算法

输入:

当前场景的描述性文字 Q

返回结果的数量 k

记忆合集 D

对优先级的重视程度 α

输出:

被选出的经验合集 *SelectedExperiences*

1: 初始化 *SelectedExperiences* 为空集合

2: 初始化 *Scores* 为空列表, 用于存储每个经验的相似度分数

3: **for** $i = 1, 2, \dots, k$ **do**

4: 根据式(2)计算 Q 和 D_i 的相似度

5: 将 $(D_i, score_i)$ 添加到 *Scores*

6: **end for**

7: 对 *Scores* 根据 $score_i$ 降序排序

8: 选取 *Scores* 中的前 k 个经验

9: **for** $i = 1, 2, \dots, k$ **do**

10: 如果经验 D_i 是由反思模块纠正的, 则 $\varepsilon = 1$; 如果是安全行驶的经验, 则 $\varepsilon = 0$

11: 根据式(3)计算每个经验的优先级

12: 根据式(4)计算每个经验被选取的概率 $P(i)$

13: **end for**

14: 根据 $P(i)$ 选取经验, 添加到 *SelectedExperiences*

15: 返回 *SelectedExperiences*

场景相似度计算公式如式(1)所示。

$$\text{similarity_search_with_score}(Q, k) = \{(D_i, score_i)\}_{i=1}^k \quad (1)$$

其中, Q 表示当前场景描述性文字, k 表示返回结果的数量, D 表示记忆合集, D_i 表示第 i 个最相似的经验, $score_i$ 表示嵌入向量与记忆模块中经验的相似度分数, 计算公式如式(2)所示。

$$score_i = \frac{\sum_{j=1}^n q_j \times d_{i,j}}{\sqrt{\sum_{j=1}^n q_j^2} \times \sqrt{\sum_{j=1}^n d_{i,j}^2}} \quad (2)$$

其中, q 表示当前场景描述性文字 Q 的向量, d_i 表示第 i 个最相似的经验 D_i 的向量。此余弦相似度公式衡量了当前场景描述性文字向量 q 和记忆模块中经验向量 d_i 之间的相似度。结果范围为 $[-1, 1]$, 其中 1 表示完全相同的向量, 0 表示无相关性, -1 表示完全相反的向量。在相似性搜索中, 本文期望寻找到与当前场景的嵌入向量相似度接近 1 的历史经验。

此外, 重要经验回放算法中会根据每条经验本身的价值赋予不同的优先级, 价值越高优先级越高, 比如记忆模块中从未出现过的经验将被赋予高价值, 重复出现的

经验将被赋予低价值; 做出错误决策发生碰撞后被纠正的经验被赋予高价值, 安全行驶的经验将被赋予低价值。优先级越高被选取出来的可能性就越大。优先级的计算公式如式(3)所示。

$$p_i = \frac{1}{score_i} + \varepsilon \quad (3)$$

其中, p_i 表示第 i 个经验的优先级, ε 有两种取值, 当经验是由反思模块中纠正后的, 则 $\varepsilon = 1$; 当经验是安全行驶的经验, 则 $\varepsilon = 0$ 。在提取的 k 个与当前场景相似的记忆经验中, 根据优先级计算每个相似经验被选取的概率, 公式如式(4)所示。

$$P(i) = \frac{p_i^\alpha}{\sum_k k \times p_k^\alpha} \quad (4)$$

其中, α 表示对优先级的重视程度, 在本算法中 $\alpha = 3$ 。最终, 被选出的经验将输入到提示生成器中, 与场景编码器输出的描述性文本集成在一起。

3) 生成提示

如图 2 中的“提示生成器”所示, 每个时刻的提示由 3 个关键部分组成: 系统提示、场景描述性文本信息和重要经验回放算法提取的经验。系统提示中包含了闭环驾驶任务的简要概述, 有对任务输入输出的内容和格式的介绍, 以及控制推理过程的约束。在每个决策框架中, 会根据当前的驾驶场景构建定制的提示文字。3 部分提示内容集成后输入到 LLMs 中, LLMs 随后使用这些提示进行推理并确定当前时刻的执行动作。

4) 三步思维链

先前研究中的运动规划器的一个常见弱点是其可解释性有限, 因为在过去运动规划器是从黑盒神经网络生成规划轨迹, 而没有阐明其决策背后的原因。又由于闭环驾驶任务需要一个复杂的推理过程才能做出正确的决策, 为了解决这个问题, 提出了一种专为自动驾驶设计的新思维链推理策略, 称为“三步思维链”方法, 包括推理过程和反思过程两个部分, 推理过程用于在推理模块中要求 LLMs 按照 3 个推理步骤逐步推理, 直至最终做出决策; 反思过程用于在反思模块中要求 LLMs 按照 3 个反思步骤逐步反思, 直至找到出错原因并修改错误。在推理模块中, “三步思维链”方法包括的 3 个步骤为: 寻找可能产生影响的关键车辆、推理车辆之间的影响情况和做出决策。

首先, 根据从环境中获取的观测信息, LLMs 需要识别可能会影响 EV 驾驶状态的其他物体, 如周围车辆等; 其次, 结合自身可选择执行的动作、周围车辆的运动状态, LLMs 应推断出周围的关键车辆可能在何时、何地以及如何影响 EV。最后, 在前述分析的基础上, LLMs 需要结合人为设置的驾驶意图给出最终的驾驶动作决策。以上三步推理框架为运动规划提供了更有条理的方法, 可

以很好地应对驾驶场景中固有的复杂性和可变性,也可以避免 LLMs 直接进行决策可能会产生幻觉的现象,并确保整个决策规划过程更加透明。

以上是 LLMs-Driver 推理模块中的全部组成,通过不断重复上述 4 个步骤,可以实现利用 LLMs 完成自动驾驶闭环系统中的推理决策任务。

1.3 记忆模块

LLMs 虽然具有强大的信息处理能力,但若没有先验知识,LLMs 在处理复杂的闭环驾驶任务时便难以进行精确推理。因此,在 LLMs-Driver 中设计了记忆模块,以增强其在自动驾驶任务中的表现。在记忆模块中,存储了过去驾驶场景中的经验,包括了决策提示、场景描述、推理过程和其他有价值的信息,其中最为重要的信息是场景描述和推理过程。场景描述详细记录了当时环境的具体情况,是在推理的过程中能够从记忆模块中成功检索到相似记忆的关键。推理过程则记录了应对这种情况的相应方法,是记忆模块中最有价值的部分。记忆模块的构建过程分为 3 个阶段:初始化、记忆检索和记忆存储。

1) 初始化

记忆模块的初始化是在 LLMs-Driver 开始训练之前,根据人为的驾驶经验,选择几个典型的驾驶场景,手动的对这些场景进行描述并完成正确的推理和决策过程,以形成初始的记忆。这些记忆将给 LLMs 带来一系列的先验知识,指导 LLMs 学习正确的驾驶决策过程。

2) 记忆检索

此过程与推理模块中的“利用重要经验回放算法从记忆模块中提取经验”是相同的。在推理的过程中,将当前驾驶场景的描述性文本信息嵌入到一个向量中,利用场景相似度搜索前 k 个与该场景相似的记忆经验,再根据优先级计算这 k 个经验的采样概率,最后按照概率分布检索出记忆模块中最有价值的经验。检索出的经验与当前场景描述一起输入提示生成器中,在提示生成器完成集成后,帮助 LLMs 做出当前情况下准确的推理和决策。

3) 记忆存储

在每个回合结束后,无论 LLMs 做出了正确的推理决策还是经过反思后学会了正确的推理过程,LLMs 都获得了新的驾驶经验。LLMs-Driver 将场景描述嵌入到密钥中,并将其与推理过程相配对,形成新的经验并存储到记忆模块中。当记忆模块的存储容量到达上限后,采用改进后的先进先出经验存储算法,对长时间未使用的经验进行替换。改进后的先进先出经验存储算法流程图如图 4 所示。

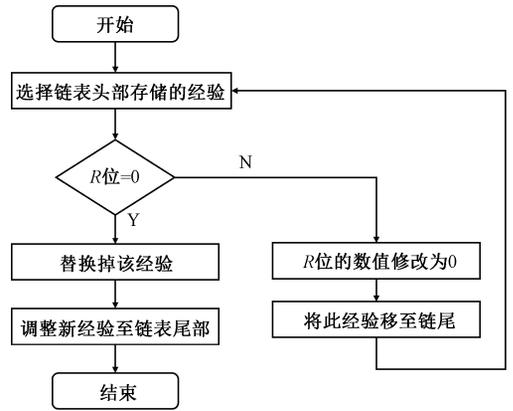


图 4 改进后的先进先出经验存储算法流程

Fig. 4 Flowchart of the improved FIFO empirical storage algorithm

具体过程为:记忆模块中的所有经验采样链表的形式进行存储。每个经验都绑定了一个标记位 R ,用来记录此经验的被检索提取的情况,标记位 R 的初始值为 0。当记忆模块中的经验被提取后,相对应的标记位 R 的数值修改为 1。当记忆模块需要新增经验但容量已达上限时,首先检查链表头部存储的经验,查看其标记位 R 的数值。若 R 的数值为 0,则表示该经验存在时间已久并未被提取使用过,直接使用新经验替换掉这条经验,并移至链表的尾部。若 R 的数值为 1,将 R 位的数值更新为 0,并将此经验移至链表的尾部进行存储。接着继续查询链表头部经验的标记位 R ,直至查找到标记位 R 为 0 的经验。

1.4 反思模块

为了使 LLMs-Driver 能够不断的积累最新的驾驶经验,掌握闭环的自动驾驶任务,引入反思模块,反思模块的结构如图 5 所示,图 5 中罗马字母表示帧数,阿拉伯数字表示优先级。在闭环驾驶任务中,将当前驾驶场景的描述和记忆模块中的过往经验作为提示生成器的输入内容,生成提示后再利用 LLMs 为每个驾驶场景生成当前时刻的决策动作。在每个回合结束后,LLMs-Driver 会得到一个决策序列,记录一个回合中每个时刻的决策动作,如图 5 中共有 30 个决策帧。LLMs-Driver 需要从每个回合的决策序列中不断反思以掌握最新的驾驶技能。LLMs-Driver 会根据决策序列中是否有冲突或者危险事件,将决策序列分为两种情况进行处理,分别为安全序列和不安全序列。

对于安全序列,即决策序列中没有任何冲突或危险事件发生,LLMs-Driver 会从决策序列中选择几个关键帧进行采样,并存入到记忆模块中。

对于不安全序列,即决策序列因有危险事件发生而终止,例如与其他车辆发生碰撞等,LLMs-Driver 会对决策序列中错误的决策进行反思,寻找原因并对决策进

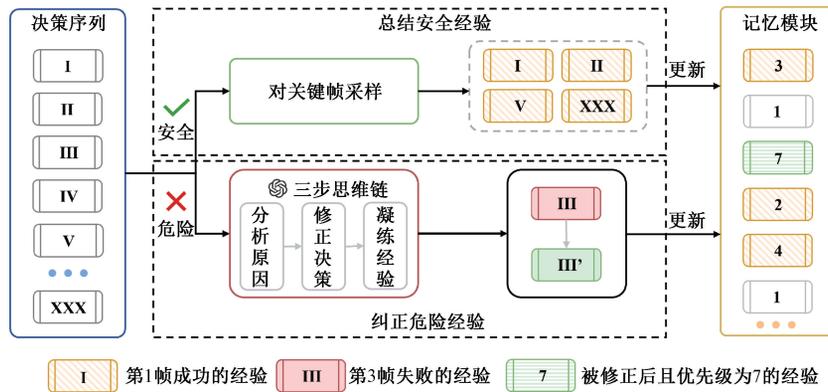


图5 反思模块结构

Fig. 5 Reflective module structure

行修正。本研究采样与推理模块中类似的“三步思维链”方法,具体包括3个步骤:分析碰撞原因、修正决策和凝练经验。首先,反思模块将发生错误决策的驾驶场景描述和原始的推理过程输入到LLMs-Driver的提示生成器中,并指示LLMs查明错误决策背后的原因。其次,反思模块要求LLMs吸取这次错误决策的经验,提出正确的策略以避免将来再次出现类似的错误。最后,LLMs-Driver把从错误中吸取的经验和修正后的决策一起存入在记忆模块中并根据式(3)计算优先级。

1.5 驾驶风格

在实际驾驶场景中,用户通常希望能够自定义自动驾驶系统的驾驶行为,以符合他们对效率和舒适度的偏好。然而,对于基于学习或基于优化的自动驾驶系统,实现这种直观和可靠的定制水平需要复杂的规则或奖励函数设计^[13]。相比之下,LLMs-Driver仅需修改向LLMs传递的驾驶意图文本即可改变驾驶风格。根据实际场景中用户驾驶的偏好,本文总结归纳了4种常见的驾驶意图,每种驾驶意图分别对应着不同的驾驶风格,具体为:

1) “安全驾驶、避免碰撞、且积极驾驶”

此驾驶风格是积极的,旨在保证安全驾驶避免发生碰撞的情况下,积极以更快的速度驾驶。这种驾驶风格下可能会出现频繁的超车、变更车道加速和减速等情况,以实现保持较高的行驶速度的目的。运动规划器需密切注意交通状况,寻找机会超越慢车,并在确保安全的前提下,尽可能减少在慢车道上的时间。尽管如此,这种风格下仍需严格遵守交通规则,避免因过于积极的驾驶行为而导致的交通事故。

2) “安全驾驶、避免碰撞、且保守驾驶”

此驾驶风格是保守的,旨在保证安全驾驶避免发生碰撞的情况下,以稳定和谨慎的速度行驶。这种风格倾向于遵循交通规则,不轻易超车,保持车辆在当前车道中行驶,避免不必要的加速、急刹车和变换车道等操作,以减少潜在的风险和事故。

3) “安全驾驶、避免碰撞、且舒适驾驶”

此驾驶风格是舒适的,保证安全驾驶避免发生碰撞的情况下,力求为乘客提供平稳的乘坐体验。这种风格避免频繁的加速和减速,以及不必要的车道变换,以减少乘客的不适感和晕车的可能性。驾驶者会保持平稳的速度,适时地加速和减速,并尽量减少急转弯和急刹车。

4) “安全驾驶、避免碰撞”

此驾驶风格是最基本的安全驾驶,主要目标是避免发生碰撞和事故。这种风格强调遵守交通规则,保持适当的车速和车距,以及对周围环境的警觉性。驾驶者会根据交通状况和道路条件调整驾驶行为,以确保自己和他人的安全,但不特别强调速度或舒适性。

2 实验及结果分析

2.1 实验设置

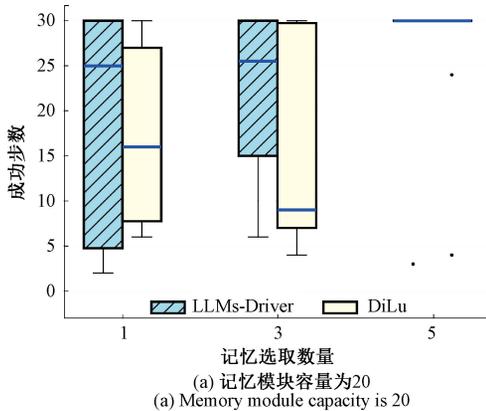
因Highway-env环境^[14]是自动驾驶决策领域中广泛使用的平台,所以本文选择在环境Highway-env上进行实验,利用高速公路场景搭建了4车道和5车道高速公路驾驶场景,来验证LLMs-Driver的性能。该场景包括了使用自主决策系统的EV和其周围的驾驶车辆。EV中配置了一个虚拟的激光雷达传感器,能够感知EV周边的行驶车辆的信息^[15-16],包括正前、左右前方、左右后方、正后的车辆。传感器的感知范围设置为150m,最大行驶速度为30m/s。Highway-env平台随机给定周围车辆的初始位置,采用智能行驶模型(intelligent driver model, IDM)和最小化变道引起的整体制动(minimizing overall braking induced by lane changes, MOBIL)模型对周围车辆进行行为决策和控制。

2.2 对比实验

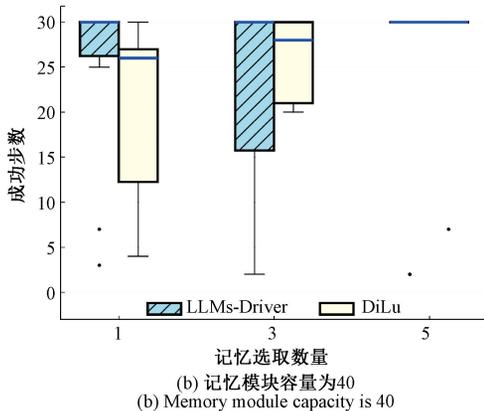
为了验证LLMs-Driver方法的有效性和泛化性,将从改变记忆模块容量、选取记忆数量、车道数量和车辆密度

方面进行实验,并与基线方法 DiLu 进行对比实验。记忆模块的容量分别设置为 20 和 40 条,推理过程中每次从记忆模块选取经验数量分别设置为 1、3 和 5 条,以验证经验积累的必要性和经验数量对性能的影响。同时,将车道数量和车道密度分别设置为 4 车道和 2 的车辆密度、5 车道和 3 的车辆密度,以验证密集交通环境下方法的性能。每种设置使用不同的随机种子重复 10 次。以成功步数(success step, SS)作为每次实验的评价指标,指没有发生碰撞的连续帧数。当 SS 为 30 时,则代表 EV 已成功完成驾驶任务。此外,考虑到箱线图能够充分反应数据的分布情况、离散程度和异常值,故选择箱线图来呈现实验结果。

关于验证 LLMs-Driver 方法有效性的实验结果如图 6 所示。



(a) 记忆模块容量为 20
(a) Memory module capacity is 20



(b) 记忆模块容量为 40
(b) Memory module capacity is 40

图 6 验证 LLMs-Driver 方法有效性实验结果

Fig. 6 Experimental results to evaluate the validity of the LLMs-Driver approach

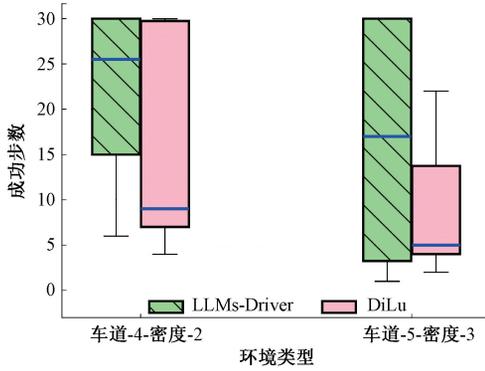
在记忆模块容量为 20 条经验的时候,LLMs-Driver 在所有分类情况下的实验中的 SS 中值均在 25 及以上,说明在大多数情况下 LLMs-Driver 均能有效学习到闭环驾驶任务的策略,结果如图 6(a) 所示。同时,结果表明 LLMs-Driver 的性能会随着选取记忆数量的增加而提高。具体来说,在每个推理环节提取数量为 1 和 3 的实验中

SS 的中值均为 25,当提取记忆数量增加到 5 时,实验几乎成功通过了所有的测试,SS 的值为 30。这可能是由于选取更多的记忆会包含类似驾驶场景中的各种经验,当这些经验作为提示输入 LLMs 时,LLMs 可以学习和利用更为广泛的信息从而做出更合理的决策。在与基线方法 DiLu 的对比中,结果表明在所有的提取经验数量的设置下,LLMs-Driver 方法的 SS 中值均大于 DiLu 方法,比如说在提取记忆数量为 3 的场景中,本方法的 SS 中值为 26,而 DiLu 的中值仅有 9。这是由于在 LLMs-Driver 中引入了重要经验回放算法和改进后的先进先出算法,在记忆模块容量固定的情况下,提出的算法能够考虑经验本身的价值并不断更新记忆模块中的经验,从而有助于 LLMs-Driver 做出更好的决策,提高了运动规划的成功率。

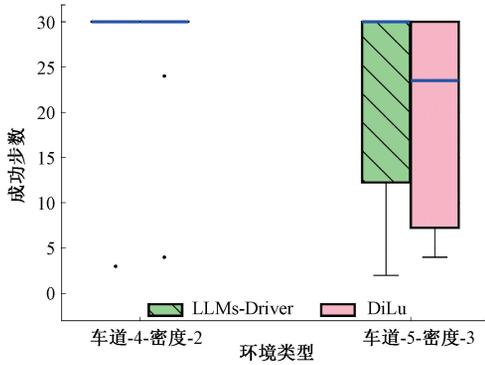
在记忆模块容量为 40 条经验的时候,LLMs-Driver 在所有分类情况下的实验中的 SS 中值均为 30,说明在所有情况下 LLMs-Driver 均能成功完成闭环驾驶任务,结果如图 6(b) 所示。而 DiLu 方法只有在选取记忆的数量为 5 时,SS 中值才能达到 30,证明了 LLMs-Driver 的学习能力和运动规划能力要强于 DiLu。对比图 6(a) 和 (b) 的性能,可以看出随着记忆模块容量的增大,可存储的经验数量增多,LLMs-Driver 在所有提取经验数量的设置下,性能均有所提高。这说明了记忆模块的容量会影响 LLMs-Driver 执行闭环驾驶任务能力,容量越大,可学习参考的适配信息越多,LLMs-Driver 性能越好。

2.3 泛化实验

关于验证 LLMs-Driver 泛化性的实验结果如图 7 所示。首先在车道数量为 4 条且车辆密度为 2 的高速公路场景中进行实验,记忆模块中获得 20 条更新后的驾驶经验,接着用更新记忆模块后的方法继续在车道数量为 5 条且车辆密度为 3 的高速公路场景中进行实验,分别测试并记录了提取记忆数量为 3 条和 5 条时的闭环驾驶性能,实验结果分别如图 7(a) 和 (b) 所示。从结果可以看出,在提取记忆数量为 3 条的情况下,将 LLMs-Driver 从车道数量为 4 条且车辆密度为 2 的高速公路场景迁移到车道数量为 5 条且车辆密度为 3 的高速公路场景后,SS 的中值从 25 降到了 17,这是符合在更为复杂的交通环境下更难取得任务成功的。在提取记忆数量为 5 条的情况下,将 LLMs-Driver 从车道数量为 4 条且车辆密度为 2 的高速公路场景迁移到车道数量为 5 条且车辆密度为 3 的高速公路场景后,SS 的中值始终保持 30 不变。这表明了在推理过程中每次提取的经验数量越多时,泛化能力会更好。与基线方法 DiLu 的实验结果进行对比,可以得出 LLMs-Driver 比 DiLu 拥有更强的泛化能力。比如说,在提取记忆数量为 5 条的情况下,迁移前后 LLMs-Driver 性能的中值并没有减小,而 DiLu 性能的中值却从 30 降到了 23。



(a) 提取记忆数量为3
(a) The number of memories extracted is 3



(b) 提取记忆数量为5
(b) The number of memories extracted is 5

图7 验证 LLMs-Driver 方法泛化性的实验结果

Fig. 7 Experimental results to verify the generalisability of the LLMs-Driver approach

2.4 可视化实验

为了比较在不同驾驶意图下 LLMs-Driver 决策风格的差异性,以及 LLMs-Driver 决策过程的可解释性,本文将可视化 LLMs-Driver 的决策过程。情景 1 中设置的驾驶风格为保守的,图中 ego 标识的车辆代表着由本文方法控制的 EV 车辆,可视化过程如图 8 所示。

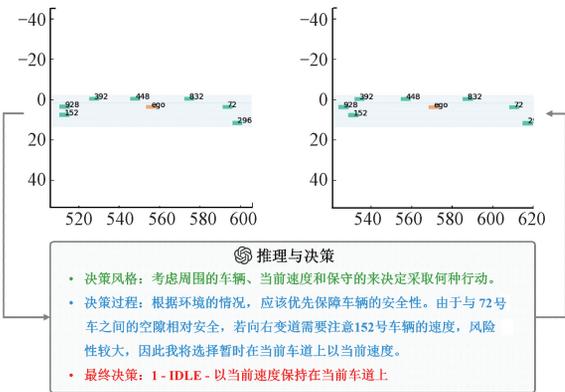


图8 情景 1 的可视化结果

Fig. 8 Visualisation results for scenario 1

首先,LLMs-Driver 对周围车辆的驾驶状态及两者之间的距离进行判断。EV 以 15.13 m/s 的速度和以 16.95 m/s 的 72 号车行驶在同一条车道上,两车相距 37.75 m,保持当前行驶状态是安全的。此外,152 号车以 17.89 m/s 的速度行驶在右侧车道,位于 EV 的后面,两车间距为 43.13 m。综合考虑两车间距和速度,EV 可以向右侧变道,但需注意 152 号车的速度。考虑到此情境下的驾驶风格是保守的,变道将伴随着更大的风险,所以最终 LLMs-Driver 做出的决策为以当前速度保持在当前车道上驾驶。这个情景有效说明了在保守的驾驶风格下 LLMs-Driver 倾向于更为稳妥风险较低的决策动作。

情景 2 中设置的驾驶风格为舒适的,期望车辆在保持安全的情况下无需过度频繁的加减速和换道,可视化过程如图 9 所示。EV 行驶在一个密集的交通环境中,以几乎相同的速度和 296 号车行驶在同一条道路上,296 号车辆在前,EV 车辆在后,768 号车辆行驶在 EV 车辆的左侧车道,960 号车辆行驶在 EV 车辆的右侧车道。经过分析,LLMs-Driver 得到向左或者向右变化车道的驾驶行为不可行,因为一方面两侧车辆的行驶速度过快且距离过近,另一方面也不符合驾驶风格为舒服的条件。因此,通过综合考虑安全性和舒适性的因素,LLMs-Driver 最终选择了保持现在的驾驶状态继续行驶。

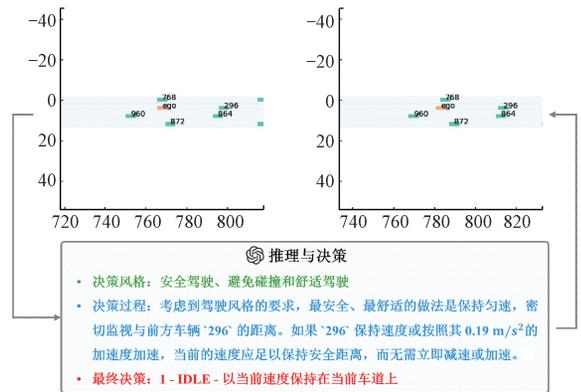


图9 情景 2 的可视化结果

Fig. 9 Visualisation results for scenario 2

情景 3 中设置的驾驶风格为积极的,可视化过程如图 10 所示。起初,EV 探索是否可以在当前车道上加速。LLMs-Driver 推理得出由于 EV 车辆速度比前方的 0 号车快,若继续加速会有撞车的风险。随后,EV 继续探索保持当前速度的可能性。由于 EV 速度比 0 车快,保持当前速度会导致两车之间的距离持续缩小,有撞车的风险。接着,EV 探索向左右换道的可行性。经推理,由于向左换道目标车道上有足够的距离和较低的速度,超车风险较低,可行性较高,并且这也满足该情境下鼓励积极超车的驾驶风格。而向右换道风险性较高,不予采纳。最后,

EV 又探讨了在当前车道减速的方案,虽然此决策是安全可行的,但是不符合此场景下的驾驶风格。故最终 EV 选择了向左变道。这也有效地说明了在被指示积极驾驶的 LLMs-Driver 会在安全保障的情况下优先做出合理的超车决定。

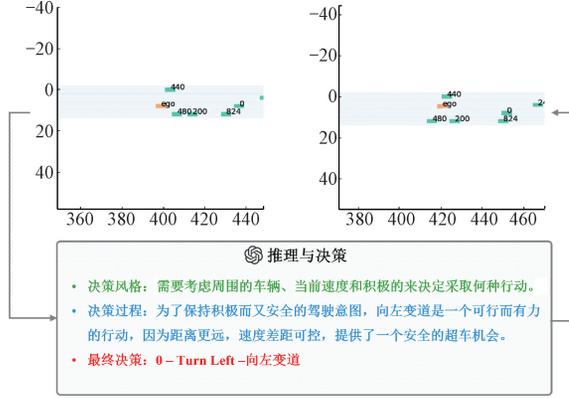


图 10 情景 3 的可视化结果

Fig. 10 Visualisation results for scenario 3

以上可视化情境分析可以直观的展示出 LLMs-Driver 思考决策的全过程,决策的动作是非常具有可解释性的,也验证了基于 LLMs 的方法相较于基于学习的方法在运动规划领域更为突出的可解释能力。同时,这也有效地说明了 LLMs-Driver 能够出色地理解用户的抽象和非直观要求,轻松实现预期驾驶风格。

此外,为了更直观的展示 LLMs-Driver 进行运动规划的全过程,本文选择了一个具有代表性的动态复杂场景进行可视化。如图 11 所示展示了在 4 车道驾驶场景中 EV 从车辆启动开始分别经过减速、连续两次变道和加速的运动过程。EV 标识的车辆表示本文算法控制的车辆。图 11 中车辆历史轨迹越密集表示车辆运动速度越慢,越稀疏表示车辆运动速度越快。在时间 $t = 0$ s 时,是环境的初始化状态。EV 和周围车辆的位置和状态由平台随机生成。在时间 $t = 1$ s 时,由于前车速度较慢,两车之间的车间距不断缩小,EV 开始减速,轨迹逐渐变得密集。在时间 $t = 2$ 和 3 s 时,EV 和前车的车间距逐渐拉大。在时间 $t = 4$ 和 5 s 时,EV 开始尝试通过变道的操作,移动到行驶速度更快的车道上。通过观察轨迹的稀疏情况可知,从左往右数第 2 车道是当前行驶最快的车道。在时间 $t = 6$ s 时,EV 完成两次变道并行驶在第 2 车道上。在时间 $t = 11$ s 时,EV 在第 2 车道上持续加速直至任务完成。以上可视化结果表明,LLMs-Driver 能够在复杂动态的驾驶场景中根据环境的具体情况进行动态决策,能够顺利完成运动规划任务,具备卓越的运动规划能力。

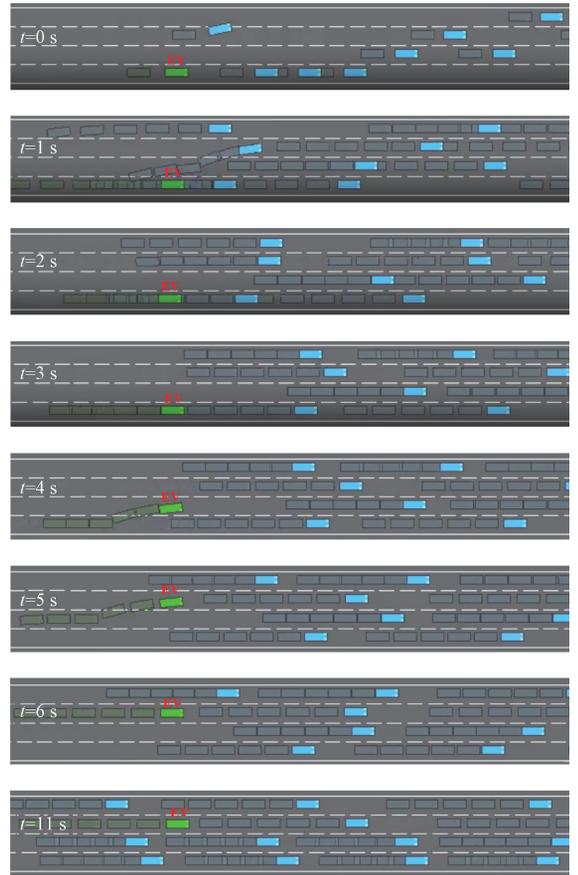


图 11 EV 运动规划任务全过程的可视化结果

Fig. 11 Visualisation results of the whole process of the EV motion planning task

2.5 消融实验

为了进一步探讨在推理模块中引入重要经验回放算法和在记忆模块中引入改进后的先进先出算法的意义,本文将选择记忆模块容量为 20 且每次提取记忆数量为 3 的实验设置进行消融实验,观察两个新算法的引入后性能的变化。实验结果如表 1 所示。

表 1 验证两个算法有效性的实验结果

Table 1 Experimental results to verify the effectiveness of the two algorithms

方法	MIN	Q1	Median	Q2	MAX
基线方法	4.0	6.8	9.0	29.3	30.0
+改进后的先进先出存储方法	4.0	11.0	23.5	30.0	30.0
+重要经验回放算法	5.0	12.5	24.0	30.0	30.0
+两个算法	6.0	13.5	25.5	30.0	30.0

基线方法为仅含有推理模块、记忆模块和反思模块 3 个模块的 DiLu 方法,并在模块容量为 20 且每次提取记

忆数量为3情况下进行实验。此外,在基线算法中陆续引入了改进后的先进先出算法和重要经验回放算法。从实验结果中可以看出,基线方法的SS中位数不到10。在分别引入了两个算法后中位数均超过了20,同时其他所有的评价指标也均取得了更高的SS。因此,在基线算法中引入改进后的先进先出算法和重要经验回放算法是合理且有效的。

3 结 论

针对基于学习的运动规划方法可解释性差的问题,提出了一个基于LLMs并具备不同驾驶风格的自动驾驶运动规划方法,称为LLMs-Driver。LLMs-Driver由3个模块组成,分别是推理模块、记忆模块和反思模块。首先,LLMs-Driver将推理和反思的过程划分为3个步骤,称为“三步思维链”方法,用于为每个思考步骤均配有详细的解释性文字,以提高LLMs-Driver的透明度和可信度。其次,在推理模块中提出了重要经验回放算法,用于在推理过程中选取过往经验的时候能够综合考虑经验的优先级和场景相似性两个因素,从而选出综合评价价值最高的经验,以提高LLMs-Driver的推理性能。此外,在记忆模块中提出了改进后的先进先出经验存储算法,当记忆模块的存储容量到达上线的时候,根据经验的过往被选取情况,保留更有价值的经验,剔除长期未使用的经验,以提高LLMs-Driver的学习性能。最后,在Highway-env仿真环境中搭建了4车道和5车道高速公路驾驶场景来验证LLMs-Driver的性能。实验结果表明,LLMs-Driver具备卓越的运动规划能力、可解释能力、泛化能力和支持自定义设置不同的驾驶风格。此外,本文仍存在一定的不足,如实验环节仅选择单一的评价指标。在未来的工作中,可考虑增加平均推理时间、决策延迟和轨迹舒适度变化率等辅助指标,以实现更为全面的性能评估。

参考文献

- [1] YURTSEVER E, LAMBERT J, CARBALLO A, et al. A survey of autonomous driving: Common practices and emerging technologies [J]. IEEE Access, 2020 (8): 58443-58469.
- [2] 张晓彤, 王嘉诚, 何景涛, 等. 面向不确定性环境的自动驾驶运动规划: 机遇与挑战[J]. 模式识别与人工智能, 2023, 36(1): 1-21.
ZHANG X T, WANG J CH, HE J T, et al. Autonomous driving motion planning for uncertain environments: Opportunities and challenges [J]. Pattern Recognition and Artificial Intelligence, 2023, 36(1): 1-21.
- [3] 李宁, 魏登, 曹裕捷, 等. 自动驾驶电动汽车避障控制算法[J]. 仪器仪表学报, 2021, 42(5): 199-207.
LI N, WEI D, CAO Y J, et al. Obstacle avoidance control algorithm for self-driving electric vehicles [J]. Chinese Journal of Scientific Instrument, 2021, 42(5): 199-207.
- [4] 韩晓庆, 姜媛媛, 刘延彬. 基于凸优化平滑动态窗口法的运动规划方法[J]. 电子测量与仪器学报, 2023, 37(8): 234-240.
HAN X Q, JIANG Y Y, LIU Y B. Motion planning method based on convex optimally smoothed dynamic window approach[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 234-240.
- [5] 姚彬, 赵盼, 林玲龙, 等. 基于本体的非结构化道路场景建模和行为决策的方法研究[J]. 电子测量技术, 2024, 47(22): 111-119.
YAO B, ZHAO P, LIN L L, et al. Research on unstructured road scene modeling and behavior decision-making based on ontology [J]. Electronic Measurement Technology, 2024, 47(22): 111-119.
- [6] 任佳佳, 柳寅奎, 胡学敏, 等. 面向复杂交通场景的自动驾驶运动规划模型[J]. 计算机工程与应用, 2024, 60(15): 91-100.
REN J J, LIU Y K, HU X M, et al. A motion planning model for complex traffic scenarios with autonomous driving [J]. Computer Engineering and Applications, 2024, 60(15): 91-100.
- [7] 朱冰, 贾士政, 赵健, 等. 自动驾驶车辆决策与规划研究综述[J]. 中国公路学报, 2024, 37(1): 215-240.
ZHU B, JIA SH ZH, ZHAO J, et al. A review of research on decision-making and planning for self-driving vehicles [J]. China Journal of Highway, 2024, 37(1): 215-240.
- [8] KIRAN B R, SOBH I, TALPAERT V, et al. Deep reinforcement learning for autonomous driving: A survey [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(6): 4909-4926.
- [9] LIU J B, FENG Y, JING SH C, et al. Deep reinforcement learning-based lane-changing trajectory planning for connected and automated vehicles [C]. 2023 9th International Conference on Mechanical and Electronics Engineering, 2023: 406-412.
- [10] FU D CH, LI X, WEN L CH, et al. Drive like a human: Rethinking autonomous driving with large

language models [C]. 2024 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops, 2024: 910-919.

- [11] WEN L CH, FU D CH, LI X, et al. DiLu: A knowledge-driven approach to autonomous driving with large language models [C]. International Conference on Learning Representations, 2024: 1-20.
- [12] CUI C, MA Y SH, CAO X, et al. Drive as you speak: Enabling human-like interaction with large language models in autonomous vehicles [C]. 2024 IEEE/CVF Winter Conference on Applications of Computer Vision, 2024: 902-909.
- [13] CHANG W J, TANG CH, LI CH R, et al. Editing driver character: Socially-controllable behavior generation for interactive traffic simulation [J]. IEEE Robotics and Automation Letters, 2023, 8(9): 5432-5439.
- [14] LEURENT E. Safe and efficient reinforcement learning for behavioural planning in autonomous driving [J]. Engineering, Computer Science, 2020.
- [15] 陈熙源, 周云川, 钟雨露, 等. 基于变分贝叶斯的鲁棒自适应因子图优化组合导航算法 [J]. 仪器仪表学报, 2024, 45(1): 120-129.
CHEN X Y, ZHOU Y CH, ZHONG Y L, et al. Robust adaptive factor graph optimization integrated navigation algorithm based on variational Bayesian [J]. Chinese Journal of Scientific Instrument, 2024, 45(1): 120-129.
- [16] 陈熙源, 戈明明, 姚志婷, 等. 雨雪天气下的激光雷达滤波算法研究 [J]. 仪器仪表学报, 2023, 44(7): 172-181.
CHEN X Y, GE M M, YAO ZH T, et al. Research on Lidar filtering algorithm for rainy and snowy weather [J]. Chinese Journal of Scientific Instrument, 2023, 44(7): 172-181.

作者简介



陈熙源 (通信作者), 1998 年毕业于东南大学获得博士学位, 现为东南大学教授, 主要研究方向为惯性技术、软件接收机及多源组合导航。

E-mail: chxiyuan@seu.edu.cn

Chen Xiyuan (Corresponding author) received his Ph.D. degree from Southeast University (SEU) in 1998. He is currently a professor at SEU. His main research interests include inertial technologies, software defined receiver and multi-source integrated navigation.



刘炜焱, 2021 年于重庆邮电大学获得学士学位, 现为东南大学硕士研究生, 主要研究方向为自动驾驶决策规划。

E-mail: 973661447@qq.com

Liu Weiyuan received her B.Sc. degree from Chongqing University of Posts and Telecommunications in 2021. She is currently a master student at Southeast University. Her main research interest includes autonomous driving perception and decision planning.



聂姝涵, 2023 年于东南大学获得学士学位, 现为东南大学硕士研究生, 主要研究方向为自动驾驶轨迹预测。

E-mail: 1016994079@qq.com

Nie Shuhan received her B.Sc. degree from Southeast University in 2023. She is currently a master student at Southeast University. Her main research interest is autonomous trajectory prediction.



经伟铭, 2023 年于三峡大学获得学士学位, 现为东南大学硕士研究生, 主要研究方向为多模态融合感知。

E-mail: weiming0120@163.com

Jing Weiming received his B.Sc. degree from Three Gorges University in 2023. He is currently a master student at Southeast University. His main research interest is multimodal fusion perception.