

doi: 10.19562/j.chinasae.qcgc.2024.ep.003

考虑主观认知的自动驾驶汽车序贯博弈类人决策*

朱冰, 贾士政, 赵健, 韩嘉懿, 张培兴, 宋东鉴, 陈志成

(吉林大学, 汽车底盘集成与仿生全国重点实验室, 长春 130025)

[摘要] 非受控交叉路口是高动态性、强交互性的决策场景, 在其中如何让自动驾驶汽车做出与熟练驾驶人类类似的安全合理决策并顺利通过路口, 是一项极具挑战性的任务。本文充分考虑认知与决策过程的主体主观属性, 提出了一种基于序贯博弈的自动驾驶汽车交互式类人决策方法。首先, 从通行时效、空间裕度、驾乘体验和行驶安全多个维度深度挖掘了多目标驾驶诱因, 进而建立了个性化和类人化驾驶特质嵌入的博弈决策模型, 并匹配不同驾驶模式类型的驾乘人员群组; 在此基础上, 提出了模仿人类逻辑的自视角决策方案和序贯优先度的概念, 以实现滚动逐阶博弈决策的序贯模式自演化; 最后, 通过多组对比实验验证了所提方法的有效性。结果表明, 本文提出的交互式类人决策方法能够以持续交互的方式化解潜在冲突并处理安全决策问题, 同时提升了自动驾驶汽车个性决策的自然化和类人化效果。

关键词: 自动驾驶汽车; 类人决策; 序贯博弈; 主观认知; 非受控交叉路口

Human-Like Decision-Making Based on Sequential Games for Automated Vehicles Considering Subjective Cognition

Zhu Bing, Jia Shizheng, Zhao Jian, Han Jiayi, Zhang Peixing,
Song Dongjian & Chen Zhicheng

Jilin University, National Key Laboratory of Automotive Chassis Integration and Bionics, Changchun 130025

[Abstract] Uncontrolled intersections are highly dynamic and strongly interactive decision-making scenarios, in which it is a challenging task to enable automated vehicles to make safe and reasonable decisions similar to skilled drivers and pass through the intersections successfully. The subjective attributes of ontology in cognition and decision-making process are fully considered in this paper, and an interactive human-like decision-making method based on sequential games for automated vehicles is proposed. Firstly, the multi-objective driving triggers are deeply explored from multiple dimensions such as traffic efficiency, space margin, ride experience, and driving safety. Further, a game decision-making model is established, which is embedded with personalized and human-like driving characteristics and can match driver and passenger groups with different driving modes and types. On this basis, the concept of sequential priority and the self-perspective decision-making scheme that imitates human logic are proposed to realize self-evolution of sequential patterns of rolling stage game decision-making. Finally, the effectiveness of the proposed method is verified through multiple sets of comparative experiments. The results show that the interactive human-like decision-making method proposed in this paper can resolve potential conflicts and deal with safety decision-making problems in a continuous and interactive manner, while improving the naturalized and human-like effect of personalized decision-making of automated vehicles.

Keywords: automated vehicles; human-like decision-making; sequential game; subjective cognition; uncontrolled intersections

* 国家自然科学基金(52172386, 523B2105)和吉林省自然科学基金(20240101116JC, 20240101121JC)资助。

原稿收到日期为 2024 年 05 月 30 日, 修改稿收到日期为 2024 年 07 月 30 日。

通信作者: 赵健, 教授, 博士研究生导师, 工学博士, E-mail: zhaojian@jlu.edu.cn。

前言

自动驾驶汽车是未来汽车科技发展的前沿和趋势^[1]。在自动驾驶汽车技术体系中,决策起到类似于人类驾驶人大脑神经中枢系统的核心作用,类人决策是使自动驾驶汽车准确成功地完成各种驾驶任务以及自然顺畅地融入复杂动态交通环境的重点和关键所在^[2]。非受控交叉路口是典型的危险场景,车辆在非受控路口的驾驶行为具有极强的交互性、强依赖于所处的交通态势。理解人类驾驶人驾驶行为的内在驱动诱因^[3],如何使自动驾驶汽车以符合驾驶人交互决策逻辑的方式自主做出决策,并在与周围车辆的持续交互中有效化解潜在交通冲突并实现非受控交叉路口安全通行,已成为当前研究的重要领域。

非受控交叉路口通行是一个持续交互的复杂过程,具有很高的动态性且对决策安全提出了极为严苛的要求。不理性的盲目驾驶行为,以及过于保守的驾驶行为均存在一定的弊端。当前,许多方法已经被应用于自动驾驶汽车在非受控交叉路口场景中自主决策和化解冲突^[4]。Noh^[5]基于贝叶斯网络设计了一种自动驾驶汽车路口决策框架,利用数字地图预测未来路径并辨识潜在风险。面对违规行驶的车辆,也能够化解冲突,避免了决策在安全方面过于保守。时间调度方法也用于化解冲突问题^[6-7]。V2V和V2I通信技术的发展可使车辆之间相互协同,并与周围的基础设施共享信息,实现合作决策,从而减少路口冲突^[8]。尽管基于通信的协同驾驶在提高路口安全性和通行高效性方面效果显著,然而它依赖于完善的基础设置部署,这在短时间内还难以全面实现。Wang等^[9]提出了一种路口冲突预测方法,涉及“最小未来距离”指标,并设计了两级动态阈值进行预警。尽管当前研究在不同的方面提升了路口通行决策效果,但在这种复杂的决策场景中如何让自动驾驶汽车像熟练的驾驶人那样做出安全合理的决策并准确顺利地通过路口,同时兼顾行驶安全性、驾乘舒适性等多目标需求,仍然存在着考虑不周的不足,值得进一步深入研究。

自动驾驶汽车决策研究可总结为学习式决策^[10-12]、反应式决策^[13-15]、端到端决策^[16-18]和交互式决策。真实世界中的驾驶环境是多智能体混合的复杂动态环境,交互是增强决策智能性的关键要素。实际上,决策是动态过程,需要动态的方法来保证驾驶的安全性和类人性。博弈论更适合于动态交互式

决策,近年来被用于智能交通系统的研究^[19]。博弈论方法主要涉及合作博弈^[20]和非合作博弈^[21]。面向混合交通环境,Hang等^[22]受大脑启发构建了驾驶模型,将动态博弈用于决策。Li等^[23]对非受控交叉路口进行参数化,建立基于成对领导者—追随者关系的动态博弈模型。Tian等^[24]基于level-k博弈论建立了非受控交叉路口决策框架。与文献^[23]中的研究相比,文献^[24]中的研究适用于兼具更多交互车辆和交叉路口数量的大规模系统。Lv等^[25]所提出的类人决策方法能适应不同的驾驶风格。在博弈均衡的优化求解方面,Yan等^[26]采用逐次平均法来迭代求解,Na等^[27]建立了人机共驾决策的4种博弈范式并推导了相应的解析解。为了提高多车系统集群决策的统一性,Jia等^[28]提出了多因素嵌入的交互式决策方法,建立了博弈模式动态切换机制。所提出的方法显著提高了决策的交互性和类人性,降低了多车博弈系统的复杂度。Liu等^[29]在所开发的three-level博弈决策框架中新颖地考虑了软交通规则,在RSS的范畴内做出安全有效的决策。博弈信息的不完备性在一些研究中也着重考虑^[30]。可见,交互式决策具有可加强决策的高交互性、高拟人度、高一体化、高协同性的优势。然而,智能体车辆模拟人类驾驶人从本体角度出发,基于主观认知展开序贯博弈决策尚未被充分研究,这是本研究的重点关注之一。

本文首先建立个性化和类人化驾驶特质嵌入的博弈决策模型;基于此,提出模仿人类逻辑的自视角决策方案和序贯优先度的概念;最后,通过多组对比实验验证所提方法的有效性。

1 问题陈述与决策系统架构

非受控交叉路口是典型的道路潜在冲突区域,如图1所示。由于缺少交通信号的引导,通行次序依靠车辆之间的交互试探和对抗争夺来实现。不理性的盲目驾驶行为会显著影响其他车辆的正常行驶,迫使其减速以保证行驶安全,进一步引起一系列的连锁反应,严重时会导致交通事故。过于保守的驾驶行为会导致车辆长时间内无法完成路口通行的任务,甚至会在路口处产生多次的反复博弈试探行为,大大降低了道路交通效率。自动驾驶汽车以及周围的车辆均应视为可以与之进行持续交互的智能体,故须研究多车之间具有持续交互处理能力的决策行为。博弈论的发展为交互决策提供了有力的技术工具。



图1 典型非受控交叉路口场景

师法人类驾驶智慧,本文提出了一种考虑主观认知的自动驾驶汽车序贯博弈类人决策方法,以解决非受控交叉路口冲突消解与安全通行问题。充分考虑决策顺序的序贯博弈是本文研究方法的基础之一。在真实的交通系统中,各个交通参与者可能具有不完全相同的目标和倾向,通常分别从本体的角度展开决策,体现为具有非合作意图的理性智能体,在与其他交通参与者的持续交互对抗中达成稳定的均衡状态,非合作博弈也是本研究重点关注的方面。图1场景的左下支路和右下支路中交错行驶的领头车辆作为本研究的对象,假设车辆的运动参数可以通过感知和通讯系统获取,本文重点研究多智能体车辆持续交互的决策过程。

本研究中决策系统的架构如图2所示。类人多

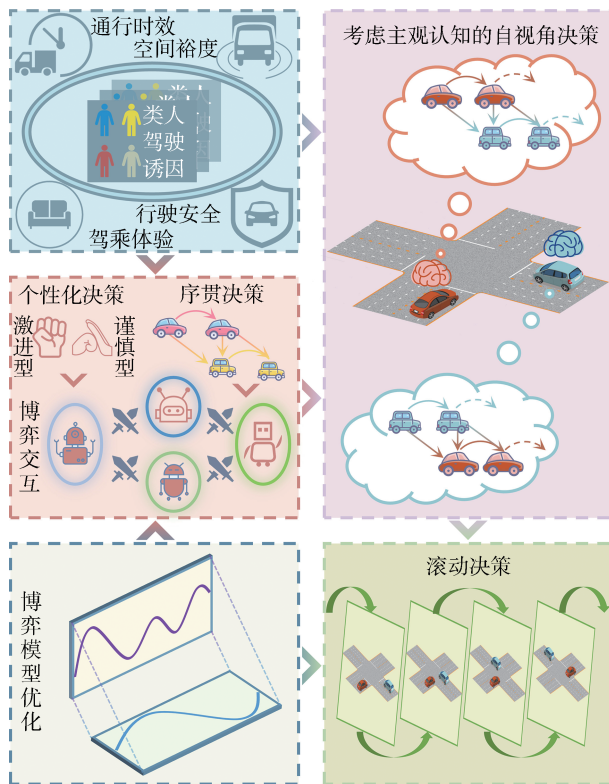


图2 决策系统架构

目标驾驶诱因建模部分是博弈交互式决策策略设计的基础,构建的决策任务充分考虑了决策的个性化特征和序贯属性,博弈模型优化部分是类人决策策略实施的支撑;以此为基础,提出了考虑主观认知的自视角决策方案,最终通过时序滚动决策完成序贯模式自演化和路口通行安全决策的任务。

2 类人决策策略

为了使自动驾驶汽车产生合理的交互决策类人逻辑,对多目标的驾驶行为内在驱动诱因进行精细化建模。博弈收益评价函数中考虑了多样化的类人驾驶需求和个性化的驾驶风格,具体为通行时效、空间裕度、驾乘体验和行驶安全。

2.1 类人多目标驾驶诱因建模

通行时效旨在提高多车系统的整体通行效率,同时使自动驾驶汽车尽可能快地化解潜在交通冲突并完成路口通行任务。在第 s 轮的阶段博弈中,通行时效收益为 \mathcal{T}_m^s 、 \mathcal{T}_n^s ,依赖于博弈双方分别从各自的博弈决策策略集中采取的任意策略以及本轮博弈开始时的瞬时运动状态。在第 s 轮的阶段博弈中,车辆 m 的通行时效收益 \mathcal{T}_m^s 和车辆 n 的通行时效收益 \mathcal{T}_n^s 如式(1)和式(2)所示。

$$\mathcal{T}_m^s = \frac{v_m(\tau_s)T - P_m(\tau_{s>s+1})}{v_m(\tau_s)} \quad (1)$$

$$\mathcal{T}_n^s = \frac{v_n(\tau_s)T - P_n(\tau_{s>s+1})}{v_n(\tau_s)} \quad (2)$$

$$v_m(\tau_s) = v_m(\tau_{s-1}) + \int_{\tau_{s-1}}^{\tau_s} a_{m,s-1}(\tau) d\tau \quad (3)$$

$$v_n(\tau_s) = v_n(\tau_{s-1}) + \int_{\tau_{s-1}}^{\tau_s} a_{n,s-1}(\tau) d\tau \quad (4)$$

$$P_m(\tau_{s>s+1}) = \int_{\tau_s}^{\tau_{s+1}} (v_m(\tau_s) + \int_{\tau_s}^{\tau_{s+1}} a_{m,s}(\tau) d\tau) d\tau \quad (5)$$

$$P_n(\tau_{s>s+1}) = \int_{\tau_s}^{\tau_{s+1}} (v_n(\tau_s) + \int_{\tau_s}^{\tau_{s+1}} a_{n,s}(\tau) d\tau) d\tau \quad (6)$$

$$\Xi(\mathcal{T}_{min}^s) = \frac{\mathcal{T}_{min}^s - \Theta_{\mathcal{T}}}{\Omega_{\mathcal{T}}} \quad (7)$$

式中: $a_{m,s}$ 、 $a_{n,s}$ 和 $a_{m,s-1}$ 、 $a_{n,s-1}$ 是车辆 m 、 n 在第 s 轮和第 $s-1$ 轮的阶段博弈中分别从各自的博弈决策策略集中采取的任意加速度策略; v_m 、 v_n 是车辆 m 、 n 的行驶速度; $T = \tau_s - \tau_{s-1} = \tau_{s+1} - \tau_s$; $P_m(\tau_{s>s+1})$ 、 $P_n(\tau_{s>s+1})$ 是车辆 m 、 n 第 s 轮博弈策略实施前后的位移, $P_m(\tau_{s>s+1}) = P_m(\tau_{s+1}) - P_m(\tau_s)$, $P_n(\tau_{s>s+1}) = P_n(\tau_{s+1}) - P_n(\tau_s)$; $\Xi(\mathcal{T}_{min}^s)$ 是针对不同量纲和维度的归一化函数; \mathcal{T}_{min}^s 是车辆 m 、 n 的通行时效收益;均值

$$\Theta_{\mathcal{T}} = \frac{1}{s} \sum_{\xi=1}^s \mathcal{T}_{m|n}^{\xi}; \text{方差 } \Omega_{\mathcal{T}} = \left(\frac{1}{s} \sum_{\xi=1}^s (\Theta_{\mathcal{T}} - \mathcal{T}_{m|n}^{\xi})^2 \right)^{0.5}.$$

空间裕度描述了驾驶人在驾驶过程中与周围车辆的距离感和对空间的占有度。在第 s 轮的阶段博弈中,空间裕度收益为 $\mathcal{M}_{m \leftrightarrow n}^s$ 和 $\mathcal{M}_{n \leftrightarrow m}^s$,依赖于博弈双方分别采取的任意策略以及双方基于此策略组合在本阶段博弈完成时的空间距离。在第 s 轮的阶段博弈中,车辆 m 的空间裕度收益 $\mathcal{M}_{m \leftrightarrow n}^s$ 和车辆 n 的空间裕度收益 $\mathcal{M}_{n \leftrightarrow m}^s$ 如式(8)所示。

$$\mathcal{M}_{m \leftrightarrow n}^s = \mathcal{M}_{n \leftrightarrow m}^s = \ln((x_m^{s+1} + y_m^{s+1})^{0.5} + 1) \quad (8)$$

$$x_m^{s+1} = (x_m(\tau_{s+1}) - x_n(\tau_{s+1}))^2 \quad (9)$$

$$y_m^{s+1} = (y_m(\tau_{s+1}) - y_n(\tau_{s+1}))^2 \quad (10)$$

$$\Xi(\mathcal{M}_{m|n}^s) = \frac{\mathcal{M}_{m|n}^s - \Theta_{\mathcal{M}}}{\Omega_{\mathcal{M}}} \quad (11)$$

式中: $(x_m(\tau_{s+1}), y_m(\tau_{s+1}))$ 和 $(x_n(\tau_{s+1}), y_n(\tau_{s+1}))$ 是车辆 m, n 完成第 s 轮博弈决策时所抵达位置的坐标,依赖于运动状态及所采取的任意加速度策略,表征为 $x_m, y_m(\tau_{s+1}) = x_m, y_m(\tau_s) + H(a_{m,s}, v_m, \tau_s, \tau_{s+1})$ 和 $x_n, y_n(\tau_{s+1}) = x_n, y_n(\tau_s) + H(a_{n,s}, v_n, \tau_s, \tau_{s+1})$,函数 H 将相关决策变量代入式(5)和式(6); $\mathcal{M}_{m|n}^s$ 是车辆 m, n 的空间裕度收益;均值 $\Theta_{\mathcal{M}} = \frac{1}{s} \sum_{\kappa=1}^s \mathcal{M}_{m|n}^{\kappa}$;方差

$$\Omega_{\mathcal{M}} = \left(\frac{1}{s} \sum_{\kappa=1}^s (\Theta_{\mathcal{M}} - \mathcal{M}_{m|n}^{\kappa})^2 \right)^{0.5}.$$

驾乘体验旨在避免频繁地加减速切换行为并惩罚由此带来的不佳驾乘体验,使车辆尽可能地平顺行驶,以提高驾乘舒适性和燃油经济性。充分考虑了决策的时序滚动特性,在第 s 轮的阶段博弈中,驾乘体验收益为 $\mathcal{R}_m^s, \mathcal{R}_n^s$,依赖于博弈双方分别采取的任意策略以及在历史决策的阶段博弈中已实施的精炼纳什均衡策略。在第 s 轮的阶段博弈中,车辆 m 的驾乘体验收益 \mathcal{R}_m^s 和车辆 n 的驾乘体验收益 \mathcal{R}_n^s 如式(12)和式(13)所示。

$$\mathcal{R}_m^s = \frac{1}{s} \left(\sum_{\chi=1}^{s-1} (\tilde{a}_m - a_{m,\chi}^*)^2 + (\tilde{a}_m - a_{m,s})^2 \right) \quad (12)$$

$$\mathcal{R}_n^s = \frac{1}{s} \left(\sum_{\gamma=1}^{s-1} (\tilde{a}_n - a_{n,\gamma}^*)^2 + (\tilde{a}_n - a_{n,s})^2 \right) \quad (13)$$

$$\tilde{a}_m = \frac{1}{s} \left(\sum_{\chi=1}^{s-1} a_{m,\chi}^* + a_{m,s} \right) \quad (14)$$

$$\tilde{a}_n = \frac{1}{s} \left(\sum_{\gamma=1}^{s-1} a_{n,\gamma}^* + a_{n,s} \right) \quad (15)$$

$$\Xi(\mathcal{R}_{m|n}^s) = \frac{\mathcal{R}_{m|n}^s - \Theta_{\mathcal{R}}}{\Omega_{\mathcal{R}}} \quad (16)$$

式中: $a_{m,\chi}^*$ 和 $a_{n,\gamma}^*$ 是车辆 m, n 分别在历史决策的第 χ

阶段和第 γ 阶段博弈中已实施的精炼纳什均衡策略, $\forall \chi, \gamma \in [1, s-1]$; $\mathcal{R}_{m|n}^s$ 是车辆 m, n 的空间裕度收益;均值 $\Theta_{\mathcal{R}} = \frac{1}{s} \sum_{\varsigma=1}^s \mathcal{R}_{m|n}^{\varsigma}$;方差 $\Omega_{\mathcal{R}} = \left(\frac{1}{s} \sum_{\varsigma=1}^s (\Theta_{\mathcal{R}} - \mathcal{R}_{m|n}^{\varsigma})^2 \right)^{0.5}$ 。

行驶安全描述了驾驶人对路口通行的安全性考量,由此衡量是否存在潜在冲突以及冲突风险的严重程度。在第 s 轮的阶段博弈中,行驶安全收益为 $\mathcal{S}_m^s, \mathcal{S}_n^s$,依赖于博弈双方分别采取的任意策略以及双方在此策略组合下的剩余通行时间差值,并在其中融入了基于主观认知的心理接受度因子。在第 s 轮的阶段博弈中,车辆 m 的行驶安全收益 \mathcal{S}_m^s 和车辆 n 的行驶安全收益 \mathcal{S}_n^s 如式(17)和式(18)所示。

$$\mathcal{S}_m^s = \frac{|T_{m, \text{RPT}}^s - T_{n, \text{RPT}}^s|}{t_{m, \text{PASC}}} \quad (17)$$

$$\mathcal{S}_n^s = \frac{|T_{n, \text{RPT}}^s - T_{m, \text{RPT}}^s|}{t_{n, \text{PASC}}} \quad (18)$$

$$T_{m, \text{RPT}}^s = \left(\frac{2l_m^s(\tau_s)}{a_{m,s}} + \left(\frac{v_m(\tau_s)}{a_{m,s}} \right)^2 \right)^{0.5} - \frac{v_m(\tau_s)}{a_{m,s}} \quad (19)$$

$$T_{n, \text{RPT}}^s = \left(\frac{2l_n^s(\tau_s)}{a_{n,s}} + \left(\frac{v_n(\tau_s)}{a_{n,s}} \right)^2 \right)^{0.5} - \frac{v_n(\tau_s)}{a_{n,s}} \quad (20)$$

$$\Xi(\mathcal{S}_{m|n}^s) = \frac{\mathcal{S}_{m|n}^s - \Theta_{\mathcal{S}}}{\Omega_{\mathcal{S}}} \quad (21)$$

式中: l_m^s, l_n^s 是车辆 m, n 在第 s 轮的阶段博弈开始时到非受控交叉路口潜在冲突区的剩余距离; $T_{m, \text{RPT}}^s, T_{n, \text{RPT}}^s$ 是车辆 m, n 基于所采取策略以及运动状态的剩余通行时间; $t_{m, \text{PASC}}, t_{n, \text{PASC}}$ 是车辆 m, n 基于主观认知的心理接受度因子; $\mathcal{S}_{m|n}^s$ 是车辆 m, n 的行驶安全收益;均值 $\Theta_{\mathcal{S}} = \frac{1}{s} \sum_{\sigma=1}^s \mathcal{S}_{m|n}^{\sigma}$;方差 $\Omega_{\mathcal{S}} = \left(\frac{1}{s} \sum_{\sigma=1}^s (\Theta_{\mathcal{S}} - \mathcal{S}_{m|n}^{\sigma})^2 \right)^{0.5}$ 。

2.2 基于序贯博弈的决策任务

序贯博弈充分考虑了决策的动态序贯性,其主要组成部分包括:参与者群、策略空间、博弈收益和均衡状态等。参与者群记为 $\mathbf{P}_{1 \dots \omega}, \mathbf{P}_{1 \dots \omega} = \{V_1, V_2, V_3, \dots, V_{\omega}\}$;策略空间记为 $\mathbf{\Pi}_{1 \dots \omega}, \mathbf{\Pi}_{1 \dots \omega} = \pi_1 \times \pi_2 \times \pi_3 \times \dots \times \pi_{\omega}$;博弈收益记为 $\mathbf{U}_{1 \dots \omega}, \mathbf{U}_{1 \dots \omega} = \{U_1(a_1, a_2 \dots a_{\omega}), \dots, U_{\omega}(a_1, a_2 \dots a_{\omega})\}$;均衡状态记为 $\mathbf{Y}_{1 \dots \omega}, \mathbf{Y}_{1 \dots \omega} = (a_1^*, a_2^*, a_3^*, \dots, a_{\omega}^*)$ 。其中,参与者群由车辆 $V_1, V_2, V_3, \dots, V_{\omega}$ 构成; $\pi_1, \pi_2, \pi_3, \dots, \pi_{\omega}$ 是参与者 $V_1, V_2, V_3, \dots, V_{\omega}$ 的博弈决策策略集, $a_1, a_2, a_3, \dots, a_{\omega}$ 是可采取的任意具体策略, $a_1 \in \pi_1, a_2 \in \pi_2, \dots, a_{\omega} \in \pi_{\omega}$; U_1, \dots, U_{ω} 是第 $1, \dots, \omega$ 层序贯博弈参与者的收益函数,共同取决于所有参与者的策略组合 $a = (a_1, a_2 \dots a_{\omega})$; $a_1^*, a_2^*, a_3^*, \dots, a_{\omega}^*$ 是每个参与者在纳什均衡状态下的最优决策策略。

对于序贯博弈 $G = \{V_1, V_2, V_3 \dots V_\omega; \pi_1, \pi_2, \pi_3, \dots \pi_\omega; U_1, U_2, U_3 \dots U_\omega\}$, 策略组合 $Y_{1 \dots \omega} = (a_1^*, a_2^*, a_3^*, \dots, a_\eta^*, \dots, a_\omega^*)$; 若对任意参与者 V_η , 当其他参与者 $P_{1 \dots \omega}^{-\eta}$ 采取策略 $Y_{1 \dots \omega}^{-\eta}$ 时, $Y_{1 \dots \omega}^{-\eta} = (a_1^*, a_2^*, a_3^*, \dots, a_{\eta-1}^*, a_{\eta+1}^*, \dots, a_\omega^*)$, a_η^* 是任意参与者 V_η 的最优策略, 即 $U_\eta(Y_{1 \dots \omega}^{-\eta}, a_\eta^* | \Phi) \geq U_\eta(Y_{1 \dots \omega}^{-\eta}, a_\eta | \Phi), \forall a_\eta \in \pi_\eta, \eta = 1, 2, 3, \dots, \omega$, 则称策略组合 $Y_{1 \dots \omega}$ 为序贯博弈 G 的纳什均衡, Φ 是由参与者群共同构成的博弈态势。基于序贯博弈的决策任务建立为

$$\begin{cases} \max_{a_{UV,s} \in \pi_{UV}, a_{LV,s} \in \pi_{LV}} U_{UV}^s(a_{UV,s}, a_{LV,s}) = \\ \quad [U_{UV}^{1,s}(a_{UV,s}, a_{LV,s}), U_{UV}^{2,s}(a_{UV,s}, a_{LV,s}), \dots, \\ \quad U_{UV}^{i,s}(a_{UV,s}, a_{LV,s})] \\ \text{s. t. } a_{LV,s} \in \operatorname{argmax}_{a_{LV,s} \in \pi_{LV}} \{U_{LV}^s(a_{UV,s}, a_{LV,s}) = \\ \quad [U_{LV}^{1,s}(a_{UV,s}, a_{LV,s}), U_{LV}^{2,s}(a_{UV,s}, a_{LV,s}), \dots, \\ \quad U_{LV}^{j,s}(a_{UV,s}, a_{LV,s})]: o_{UV}^p(a_{UV,s}, a_{LV,s}) \geq 0\} \\ \quad O_{LV}^q(a_{UV,s}, a_{LV,s}) \geq 0 \end{cases} \quad (22)$$

式中: U_{UV}^s 和 U_{LV}^s 是序贯博弈的上层先行者和下层跟随者的博弈收益函数; i 和 j 是上层和下层博弈多目标数量; o_{UV}^p 和 O_{LV}^q 是上层和下层博弈的 p 个约束函数和 q 个约束函数。

2.3 考虑主观认知的自视角决策

多智能体车辆之间的决策序贯性是序贯博弈的关键之一, 具体策略的选择依赖于决策本体的认知经验积累和相关知识储备, 初始认知信念难以完全覆盖全部信息。熟练驾驶人驾驶行为是天然的学习蓝本, 驾驶人对环境认知的本质是“主观假设-实际观察-迭代修正”的循环过程, 智能体车辆通常先从自身角度出发实施序贯博弈决策, 自然地将本体视为序贯博弈的上层先行者, 而将共同参与博弈的周围智能体车辆视为序贯博弈的下层跟随者, 周围智能体车辆亦基于其主观认知做出自视角决策。智能体车辆实时观察周围智能体车辆实际的决策策略, 并衡量与本体主观认知的偏差, 在滚动逐阶段序贯博弈决策中同步更新认知信念。考虑主观认知的自视角决策如图3所示, 滚动逐阶段序贯博弈决策时序演变过程如图4所示。

车辆的外在行驶状态和行为动作是隐式特征的具体体现, 本研究提出序贯优先度的概念用于在持续性滚动决策中实时监控并及时调整自视角决策的主观认知和序贯行动。基于加速度、速度、平均加速度、平均速度, 建立的序贯优先度为

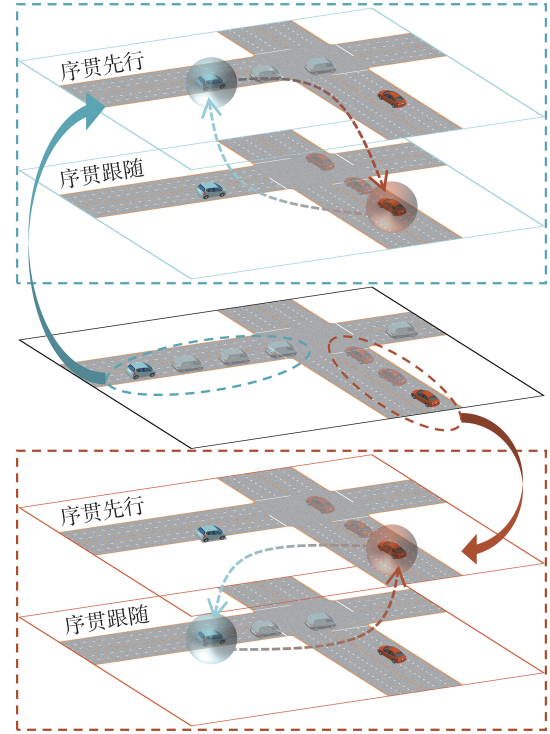


图3 考虑主观认知的自视角决策

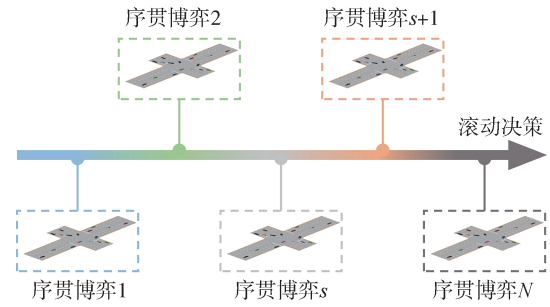


图4 滚动逐阶段序贯博弈决策时序演变过程

$$SPD_E^s = \mu_a a_E^{\theta,s} + \mu_v v_E^{\theta,s} + \mu_{\bar{a}} \bar{a}_E^{\theta,s} + \mu_{\bar{v}} \bar{v}_E^{\theta,s} \quad (23)$$

$$\bar{a}_E^s = \sum_{\delta=1}^s \frac{a_E^\delta}{s} \quad (24)$$

$$\bar{v}_E^s = \sum_{\delta=1}^s \frac{v_E^\delta}{s} \quad (25)$$

式中: $\mu_a, \mu_v, \mu_{\bar{a}}, \mu_{\bar{v}}$ 是相应的权重系数; θ 是标准化函数, $\varepsilon^\theta = (\varepsilon - \varepsilon_{\min}) / (\varepsilon_{\max} - \varepsilon_{\min})$, $\varepsilon = [a, v, \bar{a}, \bar{v}]$; E 是车辆 m, n 的集合, $E = \{m, n\}$ 。车辆 m, n 通过各自的序贯优先度 SPD_m^s, SPD_n^s 和切变阻滞阈值 $SBT_{m \rightarrow n}^{SPD}$ 在滚动决策中调整自身的序贯行动先行者或跟随者占位, 并按照序贯优先度递减的次序匹配行动占位。

2.4 类人博弈决策模型优化

单一的性能指标难以满足自动驾驶汽车驾驶需求与复杂交通系统性能要求, 通行时效、空间裕度、

驾乘体验和行驶安全等多样化多维度性能指标在优质的类人决策策略设计时是需要囊括的重点。基于建立的多目标驾驶诱因模型、序贯博弈决策任务和考虑主观认知的自视角决策逻辑,建立的多目标的类人博弈决策模型为

$$\begin{bmatrix} U_m^s \\ U_n^s \end{bmatrix} = \begin{bmatrix} \alpha_m, \beta_m, \gamma_m, \delta_m, 0, 0, 0, 0 \\ 0, 0, 0, 0, \alpha_n, \beta_n, \gamma_n, \delta_n \end{bmatrix}, \\ [\Xi(\mathcal{T}_m^s), \Xi(\mathcal{M}_m^s), \Xi(\mathcal{R}_m^s), \Xi(\mathcal{S}_m^s), \\ \Xi(\mathcal{T}_n^s), \Xi(\mathcal{M}_n^s), \Xi(\mathcal{R}_n^s), \Xi(\mathcal{S}_n^s)]^T \quad (26)$$

式中 $\alpha_m, \beta_m, \gamma_m, \delta_m, \alpha_n, \beta_n, \gamma_n, \delta_n$ 是针对个性化类人驾驶特质的权重系数,表征对多个博弈优化目标的倾向性分布。嵌入个性化特征的模型能够匹配不同类型和习惯的驾乘人员群组。谨慎型通常更注重行驶安全性,在交互博弈中常常做出让步和妥协;激进型通常表现为频繁地、持续性地激烈驾驶,通行时效是其首要关注的方面;正常型是一种中庸的类型,对多个目标有相似的重视程度。

面向类人博弈决策模型的优化任务,相应地建立了多目标协调优化约束条件。如式(22)中所示的上层和下层博弈约束条件 o_{UV}^p, O_{LV}^q ,将车辆集合E的约束条件统一记为 \mathcal{O}_E^s ,具体为

$$\mathcal{O}_E^s = [a_{E,s} - a_{\min}, a_{\max} - a_{E,s}, v_{E,s} - v_{\min}, v_{\max} - v_{E,s}, j_{\max}^2 - (\frac{a_{m,s} - a_{m,s-1}^*}{\tau_s - \tau_{s-1}})^2]^T$$

式中: a_{\max}, a_{\min} 是最大和最小加速度; v_{\max} 是道路限速条件; v_{\min} 是最小行驶速度; j_{\max} 是最大急动度;车辆集合E包含车辆 $m, n, E = \{m, n\}$ 。

3 实验结果与讨论

进行了3组典型的仿真实验以验证所提出考虑主观认知的序贯博弈类人决策方法的有效性,主要涉及博弈决策序贯性渐进稳步统一、个性化类人决策、路权保持等方面,并通过实验对比对结果进行了分析和讨论。

3.1 实验1:决策序贯性渐进稳步统一

在非受控交叉路口实验场景中,包括两辆具备主动作用和被动响应灵活机能,以及互相之间能进行持续交互决策的智能体车辆 m 和智能体车辆 n ,两车交错行驶且未来轨迹存在潜在冲突,作为图1的案例之一。在博弈系统启动前,车辆 m 和车辆 n 的初始加速度为0,初始速度为30 km/h,到潜在冲突区的初始距离为100 m,且均为谨慎型的驾驶模式

类型。可见,两车具备同等的优先度,这对于传统的基于单一行驶状态以及多维度状态融合的路权分配方法而言变得难以处理,本文提出的方法拟实现类人决策的同时,尽可能覆盖更多的棘手场景。主要实验参数如下:每辆车的加速度约束 $a_{\max} = 4 \text{ m/s}^2$, $a_{\min} = -4 \text{ m/s}^2$,道路限速条件 $v_{\max} = 90 \text{ km/h}$,急动度约束 $j_{\max} = 4 \text{ m/s}^3$,单阶段序贯博弈决策周期为0.4 s,博弈策略集的策略精度为0.2 m/s^2 。

博弈系统启动后,两车交互过程实验结果如图5和图6所示。为了与个性化类人驾驶特质紧密对应,所有实验结果采用不同的色系对其加以区分,红色系对应激进型的驾驶模式类型,蓝色系对应谨慎型的驾驶模式类型。图5(a)记录了两车序贯优先度的变化以及决策模式的切换,在 P_1 过程中,两车分别基于主观认知从自视角做出决策, P_1 过程持续了11轮阶段序贯博弈;在 P_2 过程中,车辆 m 的序贯优先度高于车辆 n 的序贯优先度且超出切变阻滞阈值,决策的序贯性收敛至车辆 m 作为上层先行者的模式, P_2 过程持续了7轮阶段序贯博弈;通过共计18轮的博弈交互,在 P_3 过程中,潜在冲突得以有效消解,两车进入博弈系统非介入模式。在起初几轮博弈交互中,两车均做出加速幅值交替增减地加速尝试,如图5(b)和图6(a-b)所示,然而随着博弈决策的时序滚动,车辆 n 逐步放缓加速直至适当减速并做出妥协;车辆 m 在获取到博弈竞争者产生让步行为的信息后,进而持续做出加速决策,如图5(b-c)所示,自然地得到路权并首先通过交叉路口,如图5(d)和图6(d)所示。图5(e)显示了交互过程的博弈收益结果。

图6分别记录了0.8、4.0、6.8、10.0 s 4个典型时刻两车交互过程的快照,图中详细记录了两车之间的决策行为和历史状态等。不同色系的矩形为相应的个性化类人驾驶特质车辆的投影,后方的尾线记录了连续3个阶段序贯博弈过程中车辆行驶的历史轨迹。红色箭头表示加速行为,绿色箭头表示减速行为。图6(a-b)处于 P_1 自视角决策阶段,图6(c)处于 P_2 序贯性统一阶段,图6(d)处于 P_3 博弈系统非介入阶段。

尽管两车产生了博弈试探和加速尝试的决策行为,但借助设计的序贯优先度和切变阻滞阈值,序贯决策模式并未反复跳变,而能够在不同模式之间实现稳步的切换。本文提出的类人决策方法有效化解了路口潜在冲突,对典型棘手场景具有一定的适应性。

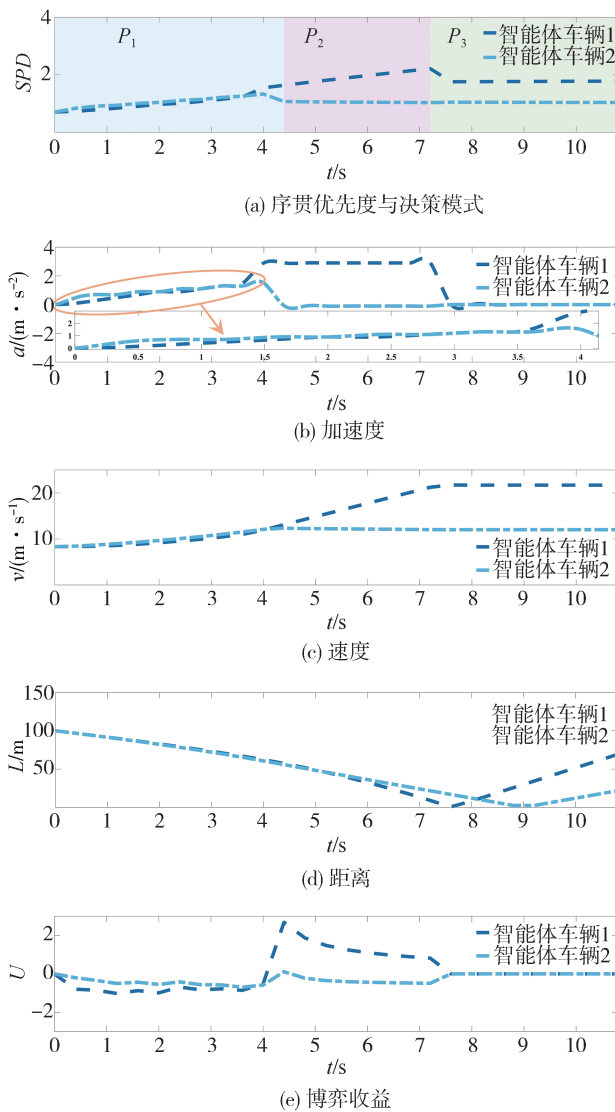


图5 实验1结果图

3.2 实验2:个性化类人决策

该实验场景作为图1的又一案例,包括车辆 m 和车辆 n 。在博弈系统启动前,车辆 m 的初始加速度为0,初始速度为30 km/h,到潜在冲突区的初始距离为100 m,且为激进型的驾驶模式类型;车辆 n 的初始加速度为0,初始速度为40 km/h,到潜在冲突区的初始距离为90 m,且为谨慎型的驾驶模式类型。综合初始行驶状态可见,车辆 n 的初始速度更大且初始距离更近,即谨慎型车辆 n 的行驶状态优于车辆 m 。在一些传统的路权分配方法中,状态占优的车辆自然地享有路权,其他车辆需要减速等待该车通过,尤其是对于该车是谨慎型的情况,其他车辆等待的时间可能会相当长,显著影响了交通系统的整体效率,且忽略了不同驾驶模式类型车辆的个性化

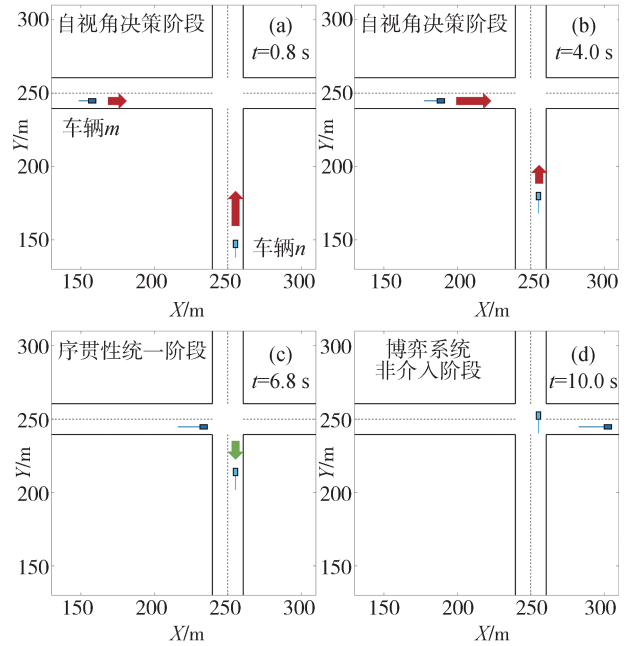


图6 实验1两车交互过程快照图

决策需求。

博弈系统启动后,基于本文所提出方法的两车交互过程实验结果如图7和图8所示,仍采用不同的色系来区分多样化的驾驶模式类型。如图7(a)所示,两车在 P_1 过程中分别基于主观认知从自视角做出决策,持续了4轮阶段序贯博弈;在 P_2 过程中,决策的序贯性收敛至车辆 n 作为下层跟随者的模式, P_2 过程持续了12轮阶段序贯博弈;通过共计16轮的博弈交互,两车在 P_3 过程中进入博弈系统非介入模式,并实现了潜在冲突消解与路口安全通行。在自视角决策 P_1 过程中,鉴于车辆 n 占优的行驶状态,其基于自身主观认知和知识经验,会做出保持自身路权优势的加速决策,尽管是谨慎型的驾驶模式类型,也并未直接减速让步,仍然产生了路权竞争和博弈试探的驾驶行为,如图7(b-c)和图8(a)所示。激进型的车辆 m 更注重通行时效性,其风格类型使其不断提升行驶速度试图优先通过路口,经过几轮的博弈交互,车辆 n 从车辆 m 的决策结果中进一步加深了对方持续竞争路权的信念,进而转为减速妥协以保证行驶安全性,与此同时,随着车辆 n 的让步,车辆 m 在滚动决策中进一步提升速度,该过程如图7(b-c)和图8(b-c)所示。图7(d)和图8(d)显示,车辆 m 在持续交互中逐步获得路权并首先通过路口。

综上,驾驶风格在一定程度上能够弥补初始状

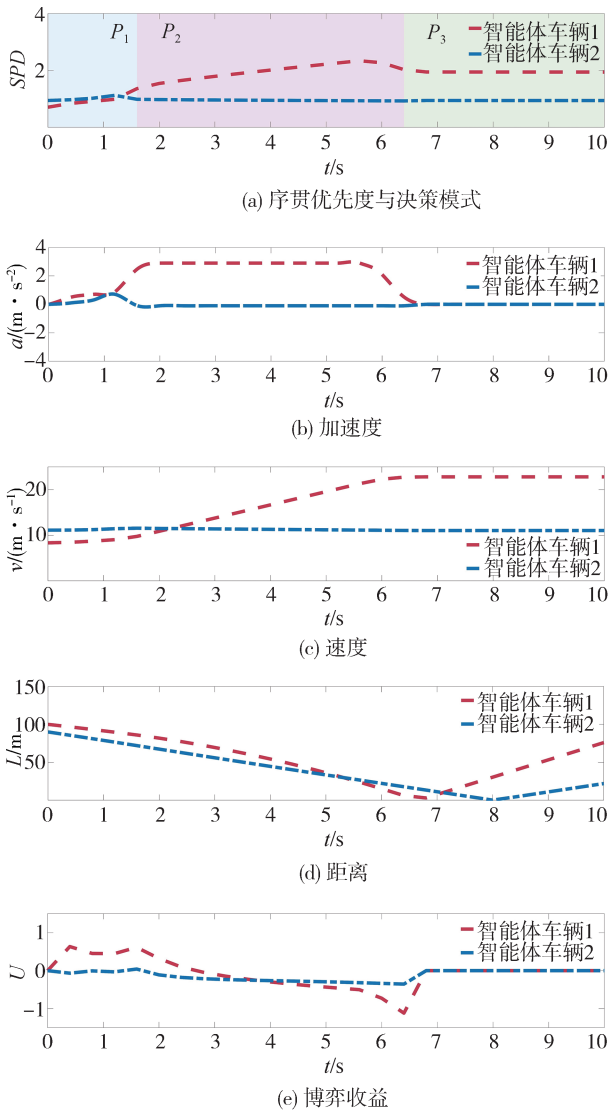


图7 实验2结果图

态的劣势,避免了减速等待谨慎型车辆而导致的低效率通行问题,本方法充分考虑了不同驾驶模式类型的自动驾驶汽车的个性化决策需求,谨慎型的车辆在经过几轮的竞争试探后才逐步妥协,本方法并未忽视智能体状态的影响及其在博弈交互决策过程中的作用。

3.3 实验3:路权保持

驾驶风格是个性化类人驾驶的重要因素之一,如上所述,智能体的行驶状态也是不可忽视的关键。该实验场景作为图1的又一案例,包括车辆 m 和车辆 n 。在博弈系统启动前,车辆 m 的初始加速度为0,初始速度为45 km/h,到潜在冲突区的初始距离为70 m,且为谨慎型的驾驶模式类型;车辆 n 的初始加速度为0,初始速度为35 km/h,到潜在冲突区的初始

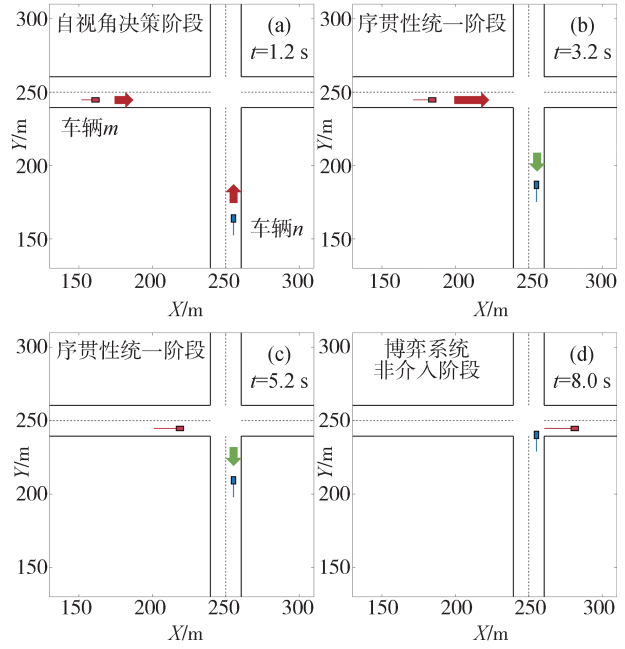


图8 实验2两车交互过程快照图

距离为100 m,且为激进型的驾驶模式类型。综合初始行驶状态可见,谨慎型车辆 m 的行驶状态显著优于车辆 n 。对于此类场景,在实际驾驶情形中,激进型车辆若产生盲目的激烈竞争行为,极易引发交通事故。

博弈系统启动后,基于本文所提出方法的两车交互过程实验结果如图9和图10所示,不同的色系仍与多样化的驾驶模式类型相匹配。如图9(a)所示, P_1 过程持续了2轮阶段序贯博弈,两车分别基于主观认知从自视角做出决策;在 P_2 过程中,决策的序贯性收敛至车辆 m 作为上层先行者的模式, P_2 过程持续了8轮阶段序贯博弈;通过共计10轮的博弈交互,有效化解了潜在冲突,两车在 P_3 过程中进入博弈系统非介入模式并先后安全通过路口。值得注意的是,如图9(b-c)和图10(a-c)所示,谨慎型车辆的行为不同于实验2中的情况,这得益于本场景中谨慎型车辆的显著状态优势,此时适当加速以迅速驶离路口相较于从高优势状态急剧减速等对方先行而言更有益于提高交通系统整体通行效率,且其路权优势得以保持。此时,车辆 m 尽管是激进型的驾驶模式类型,也并未呈现出其他实验中所显示的激烈竞争行为,而是以较小的加速动作适当轻微加速,待谨慎型车辆之后通过,在保证其通行时效性的同时也兼顾了行驶安全。该交互过程与实际的驾驶情形相符,即将通过路口的车辆的最佳策略是保持并加

强状态优势且在安全决策的前提下尽快通过路口,而不拘泥于其为哪种驾驶模式类型,正如图9(d)和图10(d)所示。

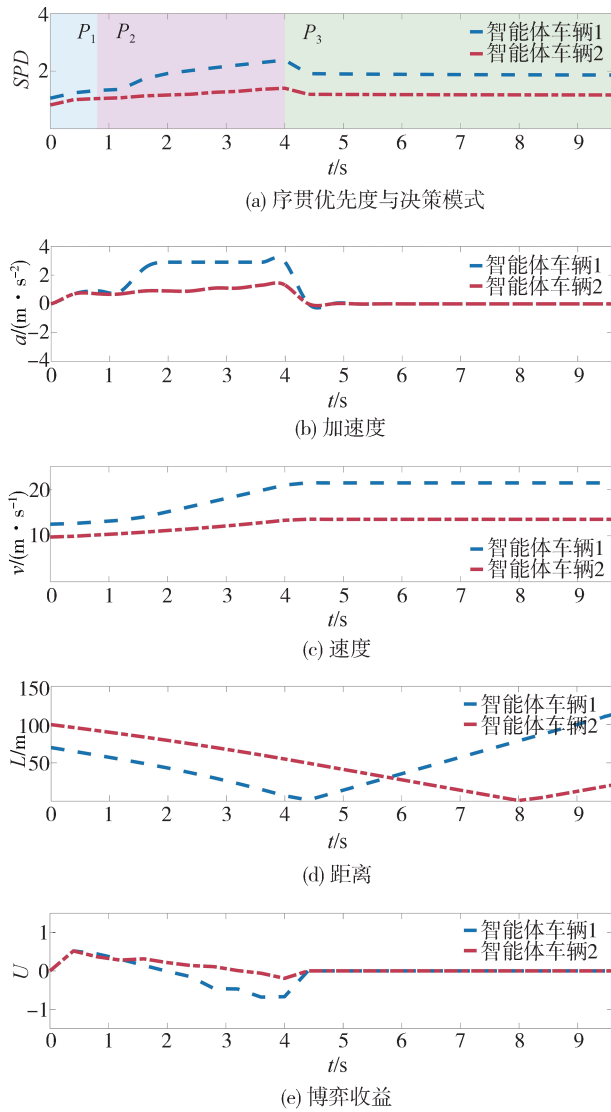


图9 实验3结果图

综合对比上述3组典型实验不难发现,博弈交互是耦合交织、动态变化的复杂过程,智能体车辆最终的行为既受到本体状态、类型的影响,也受限于所处的交通态势以及周围其他智能体的状态和类型,博弈决策结果受到多因素的共同作用。本文所提出的方法不同于传统基于规则的路权分配方法,能够以符合人类驾驶逻辑的方式与周围车辆持续交互,在滚动决策中有效化解潜在冲突并实现路口安全通行的目标,且对多种场景具有适应性,提高了自动驾驶汽车个性决策的自然化和类人化水平。

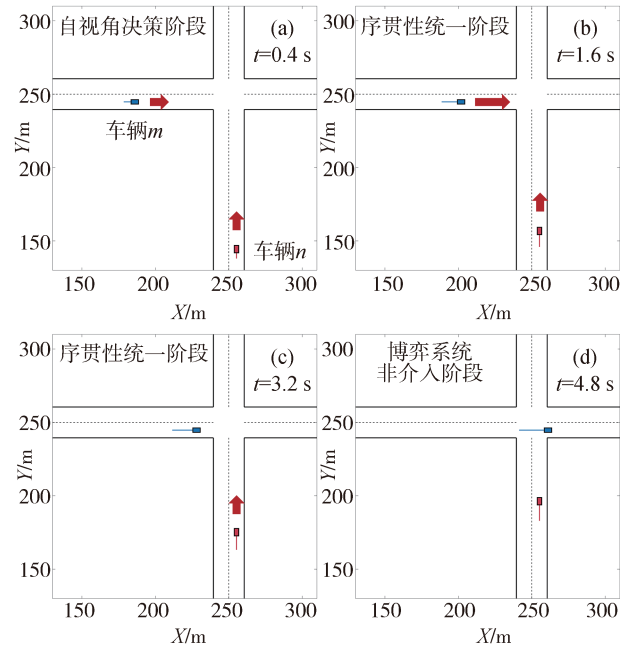


图10 实验3两车交互过程快照图

4 结论

本文面向自动驾驶汽车在动态交互的非受控交叉路口场景中的安全决策任务,提出了一种基于序贯博弈的交互式类人决策方法。本方法包括多目标驾驶诱因、个性化驾驶模式、主观认知驱动的自视角决策过程等。对比实验结果表明,交互决策的结果受到智能体本体、其他智能体以及交通态势等多重因素的耦合作用,所提出的自视角决策方案模仿了人类决策的思维逻辑,博弈试探与竞争尝试等类人行为得以体现,并实现了异构序贯决策模式的稳步切换。本方法提高了自动驾驶汽车自主决策的类人个性化程度,可自然地在智能体车辆之间进行路权动态迁移,较符合实际的人类驾驶情形。

参考文献

- [1] FENG S, SUN H, YAN X, et al. Dense reinforcement learning for safety validation of autonomous vehicles [J]. *Nature*, 2023, 615(7953): 620–627.
- [2] NOH S, AN K. Decision-making framework for automated driving in highway environments [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 19(1): 58–71.
- [3] KOLEKAR S, DE WINTER J, ABBINK D. Human-like driving behaviour emerges from a risk-based driver model [J]. *Nature Communications*, 2020, 11(1): 1–13.

- [4] HUBMANN C, SCHULZ J, BECKER M, et al. Automated driving in uncertain environments: planning with interaction and uncertain maneuver prediction[J]. *IEEE Transactions on Intelligent Vehicles*, 2018, 3(1): 5–17.
- [5] NOH S. Decision-making framework for autonomous driving at road intersections: safeguarding against collision, overly conservative behavior, and violation vehicles[J]. *IEEE Transactions on Industrial Electronics*, 2018, 66(4): 3275–3286.
- [6] AHN H, DEL VECCHIO D. Safety verification and control for collision avoidance at road intersections[J]. *IEEE Transactions on Automatic Control*, 2017, 63(3): 630–642.
- [7] DE CAMPOS G R, FALCONE P, HULT R, et al. Traffic coordination at road intersections: autonomous decision-making algorithms using model-based heuristics[J]. *IEEE Intelligent Transportation Systems Magazine*, 2017, 9(1): 8–21.
- [8] HAFNER M R, CUNNINGHAM D, CAMINITI L, et al. Cooperative collision avoidance at intersections: algorithms and experiments[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(3): 1162–1175.
- [9] WANG P, CHAN C Y. Vehicle collision prediction at intersections based on comparison of minimal distance between vehicles and dynamic thresholds[J]. *IET Intelligent Transport Systems*, 2017, 11(10): 676–684.
- [10] REN Y, ZHAN G, TANG L, et al. Improve generalization of driving policy at signalized intersections with adversarial learning[J]. *Transportation Research Part C: Emerging Technologies*, 2023, 152: 104161.
- [11] LIAO H, HE Y, WU X, et al. Reimagining multi-criterion decision making by data-driven methods based on machine learning: a literature review[J]. *Information Fusion*, 2023: 101970.
- [12] BHATTACHARYYA R, WULFE B, PHILLIPS D J, et al. Modeling human driving behavior through generative adversarial imitation learning[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 24(3): 2874–2887.
- [13] JO K, KIM J, KIM D, et al. Development of autonomous car—part II: a case study on the implementation of an autonomous driving system based on distributed architecture[J]. *IEEE Transactions on Industrial Electronics*, 2015, 62(8): 5119–5132.
- [14] HU H, JIA X, LIU K, et al. Self-adaptive traffic control model with behavior trees and reinforcement learning for AGV in industry 4.0[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(12): 7968–7979.
- [15] JIA S, LI X, LYU Y, et al. Simulation research on safety and anti-collision strategy of typical intersection based on hazard cognition[C]. 2019 Chinese Automation Congress (CAC). IEEE, 2019: 2652–2657.
- [16] LI K, XIONG H, YU D, et al. An end-to-end multi-task learning model for drivable road detection via edge refinement and geometric deformation[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(7): 8641–8651.
- [17] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. *Nature*, 2015, 518(7540): 529–533.
- [18] MIN K, KIM H, HUH K. Deep distributional reinforcement learning based high-level driving policy determination[J]. *IEEE Transactions on Intelligent Vehicles*, 2019, 4(3): 416–424.
- [19] DI X, SHI R. A survey on autonomous vehicle control in the era of mixed-autonomy: from physics-based to AI-guided driving policy learning[J]. *Transportation Research Part C: Emerging Technologies*, 2021, 125: 103008.
- [20] NA X, COLE D J. Modelling of a human driver's interaction with vehicle automated steering using cooperative game theory[J]. *IEEE/CAA Journal of Automatica Sinica*, 2019, 6(5): 1095–1107.
- [21] HANG P, LV C, HUANG C, et al. An integrated framework of decision making and motion planning for autonomous vehicles considering social behaviors[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(12): 14458–14469.
- [22] HANG P, ZHANG Y, LV C. Brain-inspired modeling and decision-making for human-like autonomous driving in mixed traffic environment[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [23] LI N, YAO Y, KOLMANOVSKY I, et al. Game-theoretic modeling of multi-vehicle interactions at uncontrolled intersections[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 23(2): 1428–1442.
- [24] TIAN R, LI N, KOLMANOVSKY I, et al. Game-theoretic modeling of traffic in unsignalized intersection network for autonomous vehicle control verification and validation[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 23(3): 2211–2226.
- [25] HANG P, LV C, HUANG C, et al. Human-like lane-change decision making for automated driving with a game theoretic approach[C]. 2020 4th CAA International Conference on Vehicular Control and Intelligence (CVCI). IEEE, 2020: 708–713.
- [26] YAN X Y, ZHOU T. Destination choice game: a spatial interaction theory on human mobility[J]. *Scientific Reports*, 2019, 9(1): 9466.
- [27] NA X, COLE D J. Game-theoretic modeling of the steering interaction between a human driver and a vehicle collision avoidance controller[J]. *IEEE Transactions on Human-Machine Systems*, 2014, 45(1): 25–38.
- [28] JIA S, ZHANG Y, LI X, et al. Interactive decision-making with switchable game modes for automated vehicles at intersections[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [29] LIU M, WAN Y, LEWIS F L, et al. A three-level game-theoretic decision-making framework for autonomous vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(11): 20298–20308.
- [30] ZHANG Q, LANGARI R, TSENG H E, et al. A game theoretic model predictive controller with aggressiveness estimation for mandatory lane change[J]. *IEEE Transactions on Intelligent Vehicles*, 2019, 5(1): 75–89.