

北京工业大学本科毕业论文结题（申优）答辩



北京工业大学
BEIJINGUNIVERSITYOFTECHNOLOGY

自动驾驶车辆利己行为建模与影响分析

答辩人：李悦（交通设备与控制工程 16030120） 导师：贺正冰 教授

城市交通学院 2020年6月12日



目录

- 1 绪论
- 2 基于Q学习的自动车利己行为建模
- 3 仿真与数值分析
- 4 论文总结



1. 绪论

- 研究背景
- 研究综述
- 研究内容
- 技术路线

研究背景

人类驾驶员的特性

- 局限性
- 异质性
- 其他

2040年：AV占比**50%**

研究 背景

自动车（AV）的优点

减少事故发生 降低人工成本
解决城市拥堵 利于环境保护

“试错”：无需人工监督，更灵活的学习方式



自动车的**智能性+不确定性**

自动车与人驾车共存的境况



强化学习等AI领域的兴起

研究综述

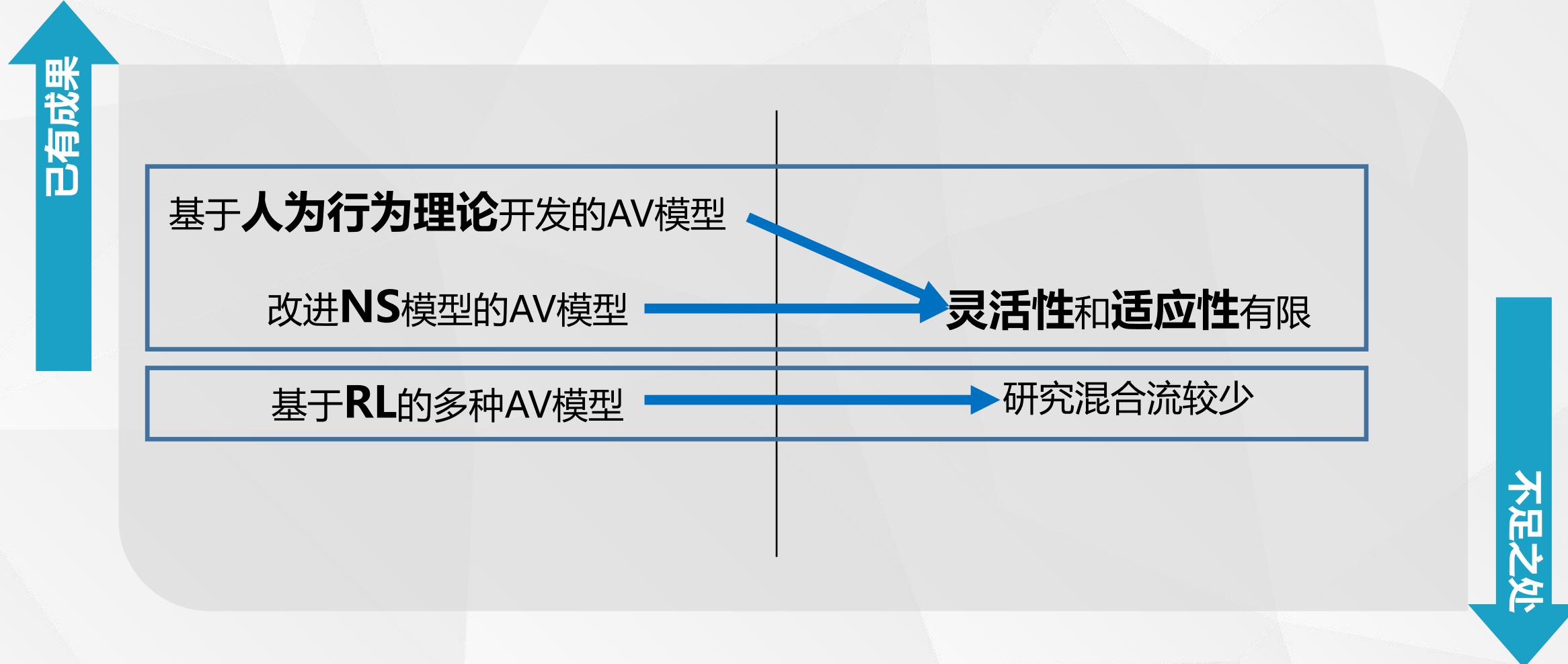
(1) 人机混合交通流

作者	论文题目	研究内容
Arnab Bose 等	Mixed manual/semi-automated traffic: a macroscopic analysis	混合流的流量变化；停走状态下的混合流
Jincai Chang等	Modeling and simulations on automated vehicles to alleviate traffic congestion	自动车对混合流通行能力的影响；最佳自动车渗透率
Danjue Chen 等	Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles	影响人机混合流通行能力的若干因素
Sina Bahrami 等	Optimal traffic management policies for mixed human and automated traffic flows	自动车数量与混合流通行能力成正相关的原因
Jiazu Zhou 等	Modeling the fundamental diagram of mixed human-driven and connected automated vehicles	不同类型的车头时距对混合流基本图的影响
Yangzexi Liu等	Characteristic Analysis of Mixed Traffic Flow of Regular and Autonomous Vehicles Using Cellular Automata	自动车及其渗透率对异构交通流动力学的影响
Bokui Chen等	A future intelligent traffic system with mixed autonomous vehicles and human-driven vehicles	自动车预见力、自动车渗透率、车辆密度、人驾车随机减速的可能性等对道路通行能力的影响

(2) 强化学习在自动车行为描述中的应用

作者	论文题目	研究内容
Meixin Zhu等	Human-like autonomous car-following model with deep reinforcement learning	基于深度强化学习的自动车跟车模型，并使用深度确定性策略梯度算法 DDPG 对模型进行优化
Jingqiu Guo 等	Merging and Diverging Impact on Mixed Traffic of Regular and Autonomous Vehicles	基于强化学习技术的网联自动车技术；出/入匝对混合流的变道行为的影响
Changxi You 等	Advanced planning for autonomous vehicles using reinforcement learning and deep inverse reinforcement learning	使用强化学习逆向强化学习技术解决交通中自动车的规划问题

研究述评

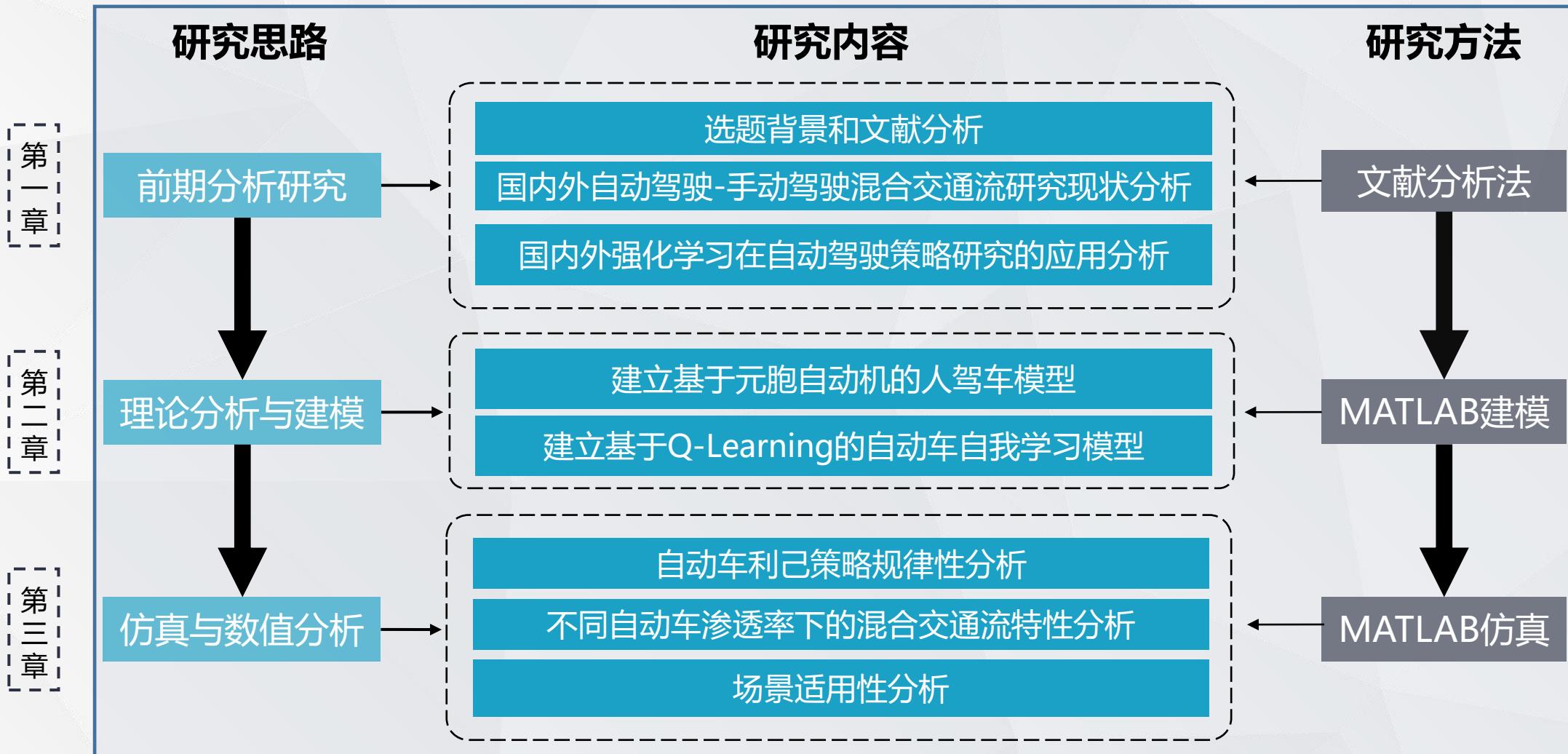


研究内容

1. 建立CA框架下的基于Q学习的自动车行驶策略**学习模型**
2. 训练不同场景下的**自动车**以获取其“利己”策略
3. 改变相关仿真参数, 分析自动车“利己”**策略对混合交通流的影响**
4. 分析并验证自动车“利己”策略**适用性**



技术路线

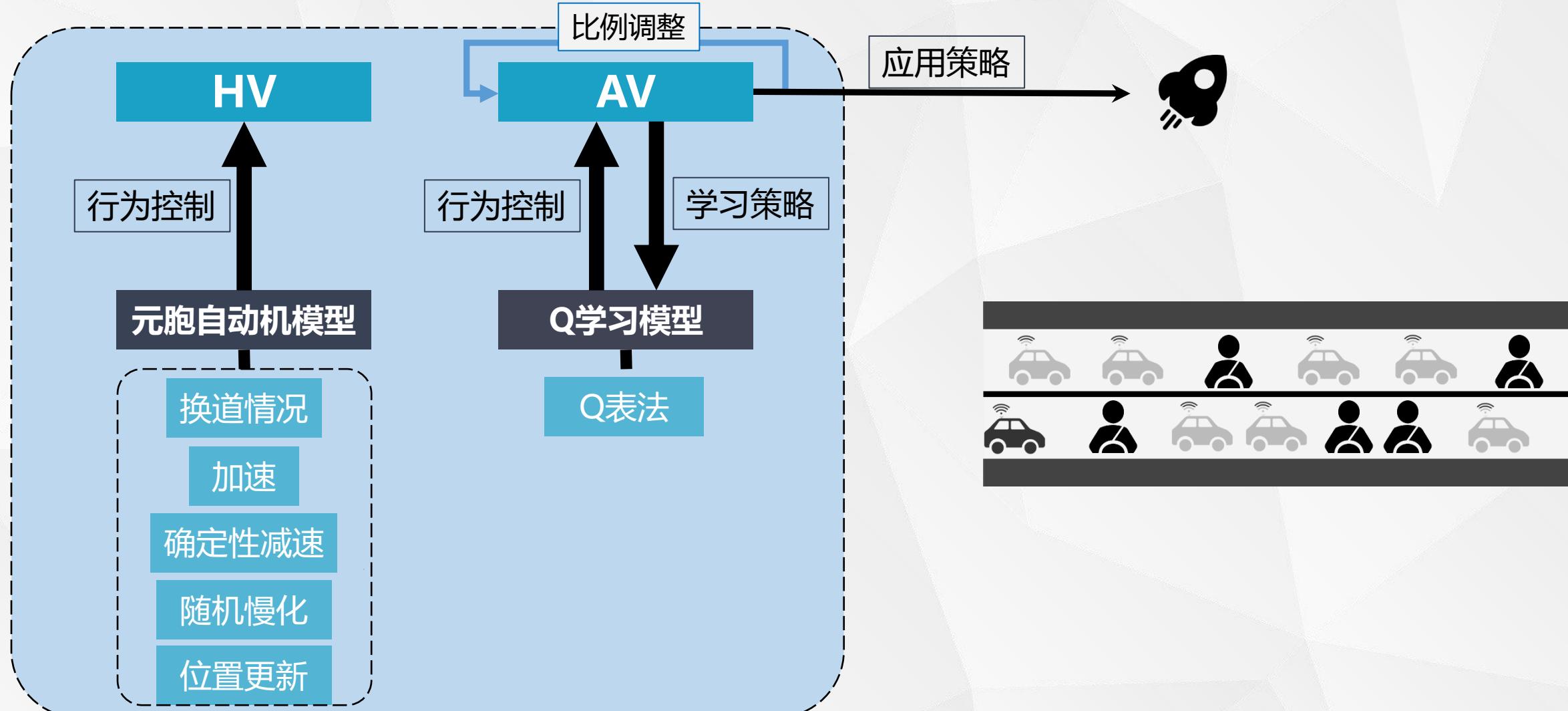




2. 基于Q学习的自动驾驶利己行为建模

- 人驾车建模
- 强化学习算法
- 自动车利己建模

训练方法 - HV+AV



人驾车建模：对称同向双车道元胞自动机(STCA)模型

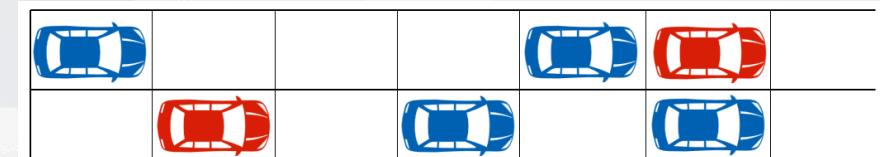
人驾车状态更新基本规则



人驾车换道规则

1. 计算换道情况
2. 加速: $v_n \rightarrow \min(v_n + 1, v_{max})$
3. 确定性减速: $v_n \rightarrow \min(v_n, d_n)$
4. 随机慢化: $v_n \rightarrow \min(v_n - 1, 0)$
5. 位置更新: $x_n \rightarrow x_n + v_n$

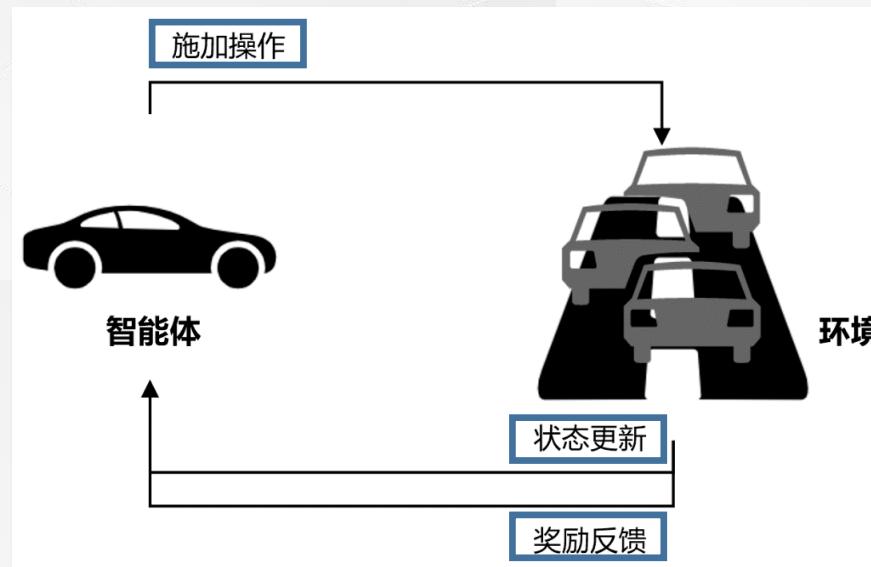
- $d_n < \min(v_n(t) + 1, v_{max})$
- $d_{n+1,other} > v_{max}$
- $d_{n,other} > d_n$
- $link_{n,other} \text{ is empty}$



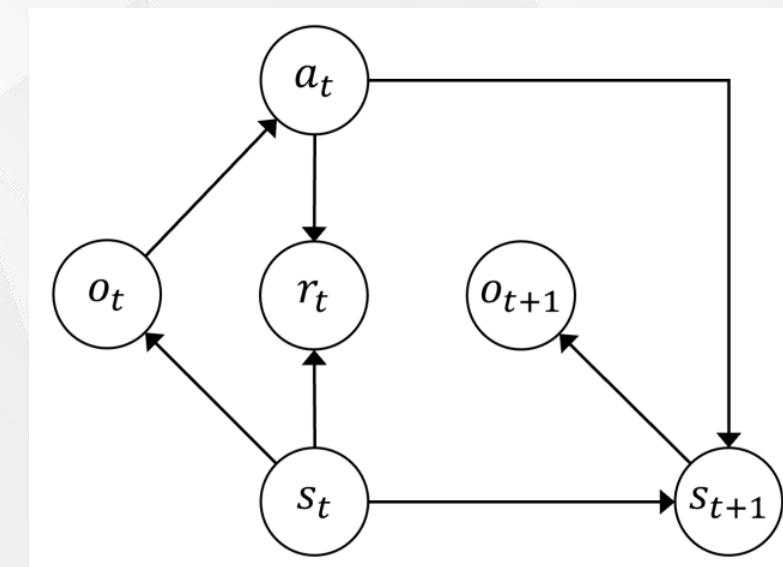
强化学习算法：马尔可夫决策过程(MDP)框架

- $p(s_{t+1}|s_t, a_t, s_{t-1}, a_{t-1}, \dots, s_0, a_0) = p(s_{t+1}|s_t, a_t)$
- $\pi^* = \arg \max_{\pi} E[\sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(s_t))]$

MDP基本作用



MDP应用于自动车建模中



强化学习算法：基于 ϵ -贪婪搜索策略的Q学习方法

Q学习步骤

① 初始化Q表：

	A_1	A_2	...
s_1	$Q(s_1, A_1)$		
s_2			
...			

② $A = [A_1, A_2 \dots] \rightarrow r_t$



③ r_t

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_t + \gamma \max_{a_{t+1} \in A} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t))$$

④ 重复②③至一个学习片段结束

⑤ 重复上述迭代过程直至运行结束

算法 1 Q-Learning using ϵ -greedy search

Input: $S, A, \alpha, \gamma, \epsilon, R$

Output: the optimal state action function $Q^*(s, a)$

```

1:  $Q \leftarrow Q_0$ 
2:  $Q(s_{final}, \cdot) \leftarrow 0$ 
3:  $Converge \leftarrow False$ 
4: while run time < time limit OR episode < episode limit do
5:    $t \leftarrow 0$ 
6:    $s_t \leftarrow s_0$ 
7:   while  $t \leq horizon\ limit$  do
8:     if  $rand() \leq \epsilon$  then
9:       Randomly choose  $a_t$  from action space;
10:      else
11:         $a_t = \arg \max_a Q(s_t, a)$ 
12:      end if
13:       $r_t \leftarrow$  reward function( $s_t$ )
14:       $s_{t+1} \leftarrow$  traffic model based on  $s_t$  and  $a_t$ 
15:      Bellman update on  $Q$ :  $Q(s_t, a_t) \leftarrow$ 
16:      
$$Q(s_t, a_t) + \alpha(r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t));$$

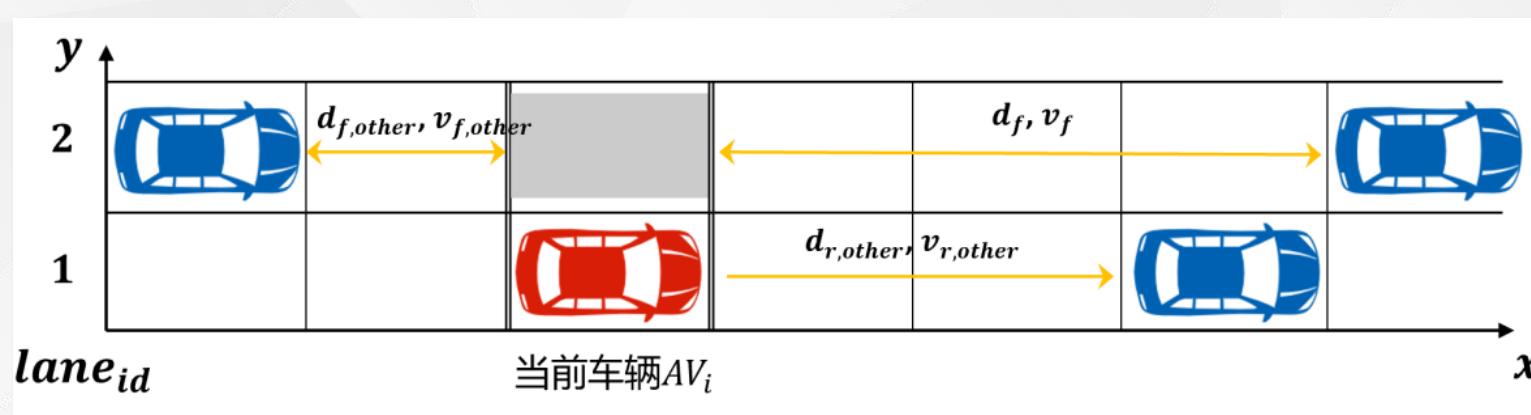
17:       $t \leftarrow t + 1$ ;
18:    end for
19:  end for
20: return  $Q^*(s, a) \leftarrow Q(s, a);$ 

```

贝尔曼方程

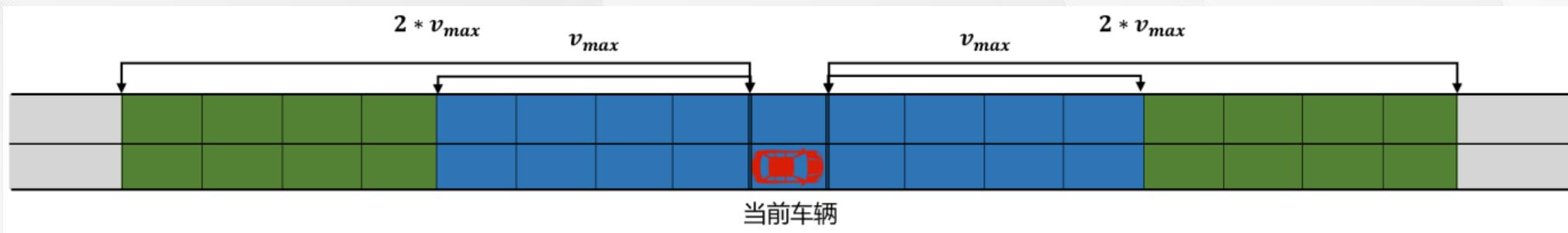
自动驾驶利己建模：状态空间 S 定义

(连续) 观察空间 $O = \{d_f, v_f, d_{f,other}, v_{f,other}, d_{r,other}, v_{r,other}, lane_{id}\}$



$$3^6 * 2 = 1458 \text{ 个状态}$$

(离散) 状态空间 $S = \{d_f, v_f, d_{f,other}, v_{f,other}, d_{r,other}, v_{r,other}, lane_{id}\}$



自动驾驶利己建模：动作空间 A 与奖励函数 R 定义

动作空间 A

一般：加速、强加速、保持、减速、强减速、换道保持、**换道加速、换道减速**



本研究：加速、强加速、保持、减速、强减速、换道保持

奖励函数 R

在仅考虑速度之差 v 的基础上，加入对强加减速的惩罚因子: a

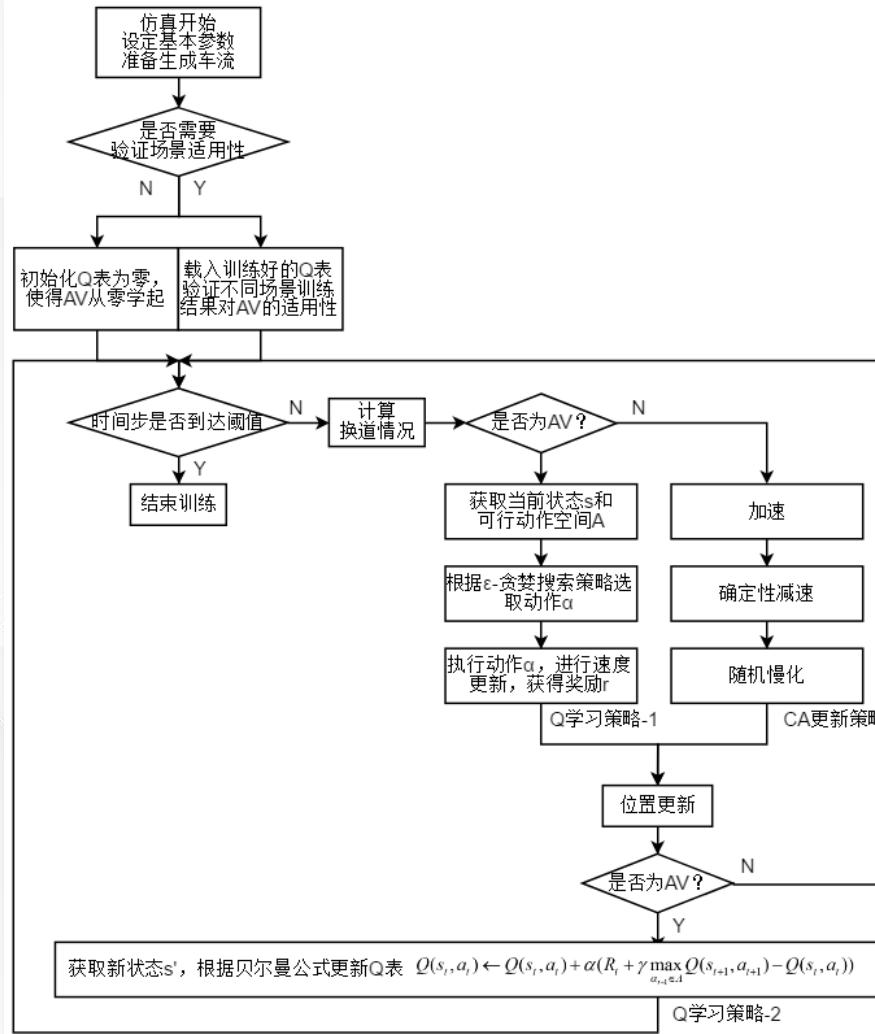
$$R = \omega_1 v + \omega_2 a$$



3. 仿真与数值分析

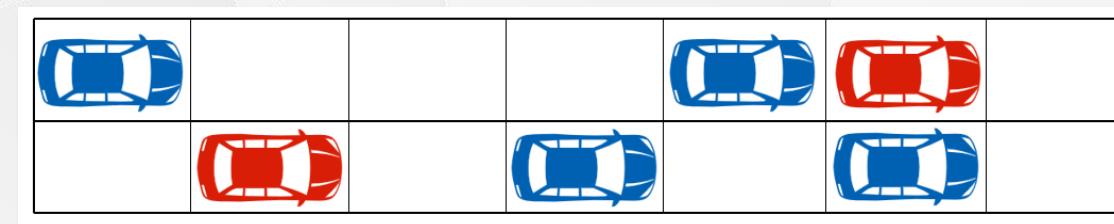
- 混合交通流特性分析
- 自动车利己策略验证与分析
- 场景适用性分析

仿真环境设置与训练



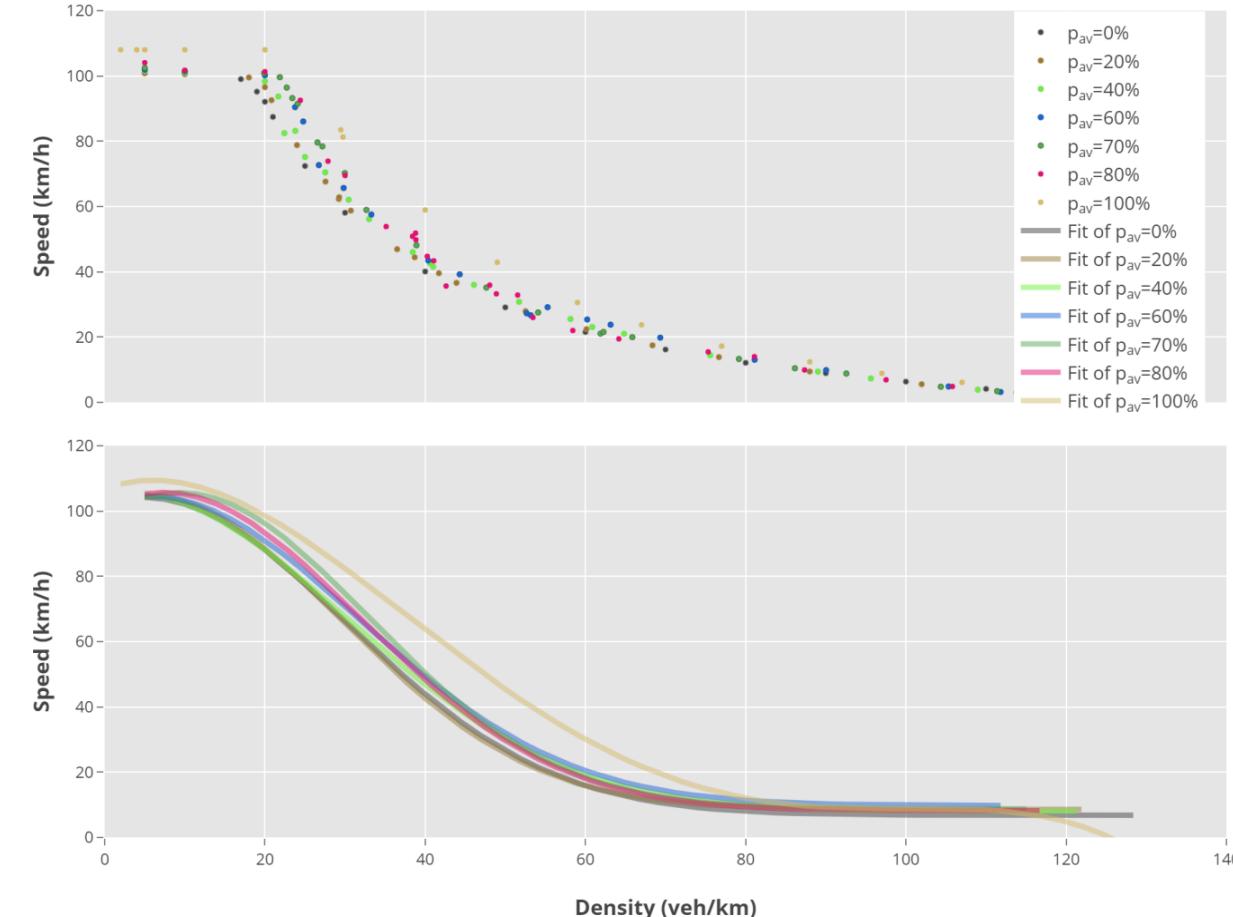
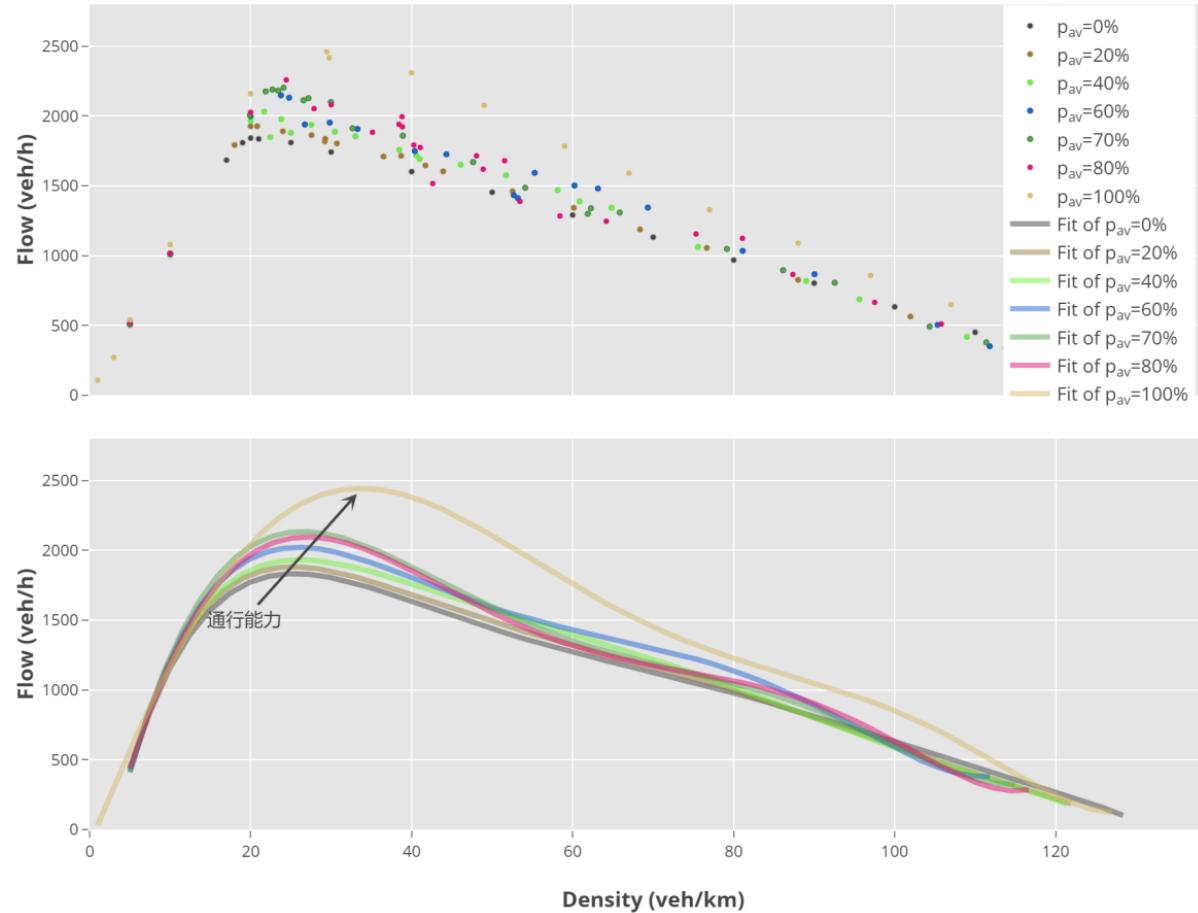
训练1：初始化Q表为零，使得AV从零学起

训练2：载入训练好的Q表，验证不同场景训练结果对AV的适用性



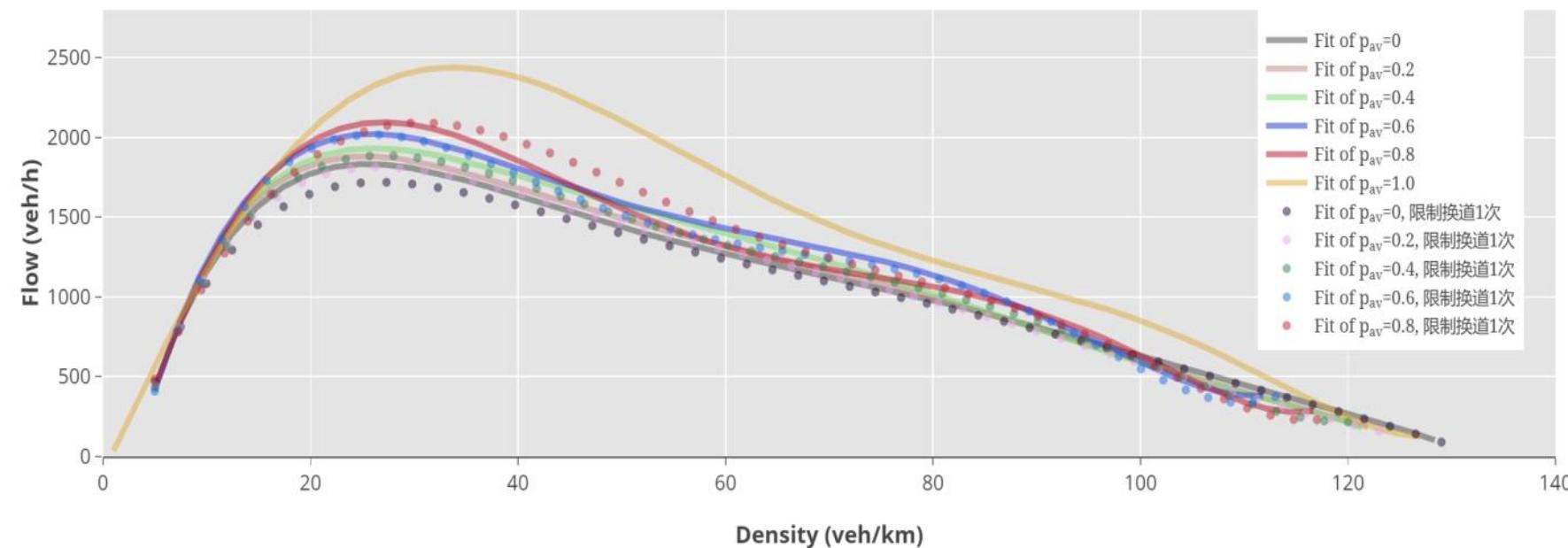
混合交通流特性分析：基本图3-1

随着自动车渗透率的提升，道路通行能力逐步上升，且混合流的平均车速也有一定的提升；同时，临界密度逐渐后移，车流稳定性提升。



混合交通流特性分析：基本图3-2

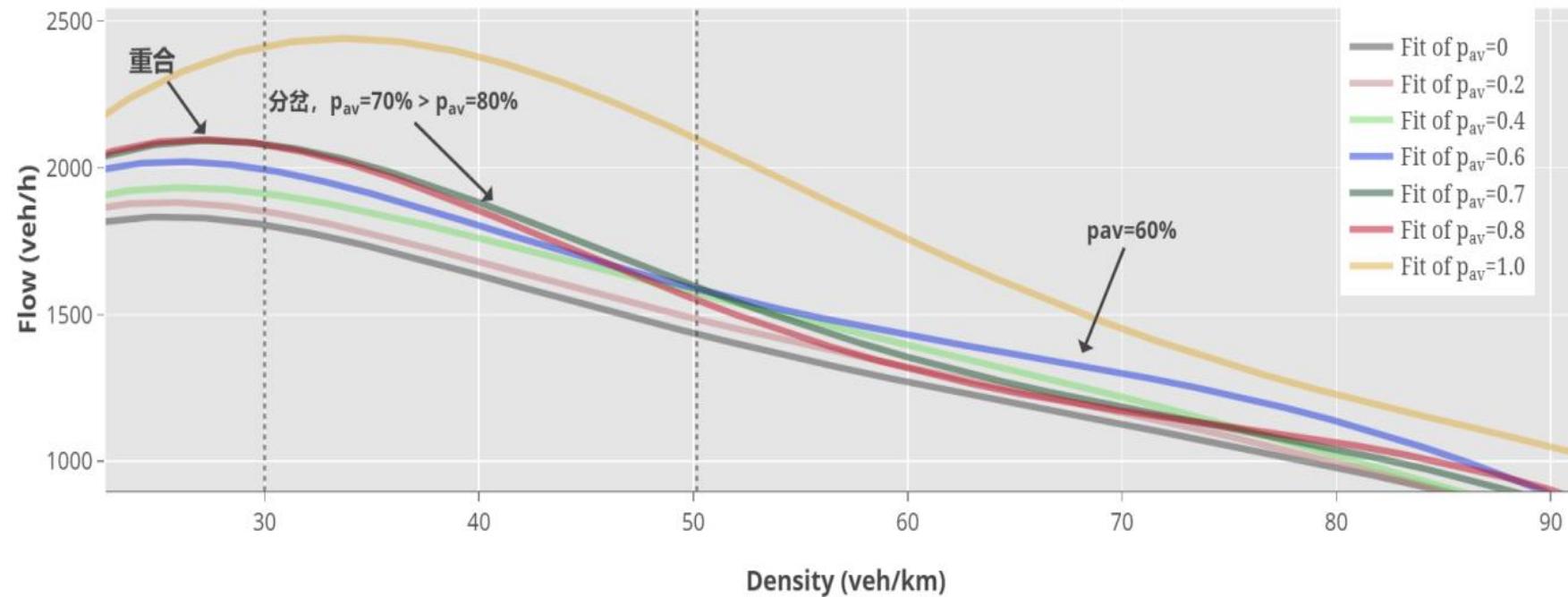
人驾车的频繁换道在**低自动车渗透率**时可以较有效地利用道路资源；但在**高自动车渗透率**时，由于人驾车并未有“预见”能力，因此频繁换道或将增加道路拥堵的可能性。



人驾车无限制换道 VS 人驾车限制换道1次

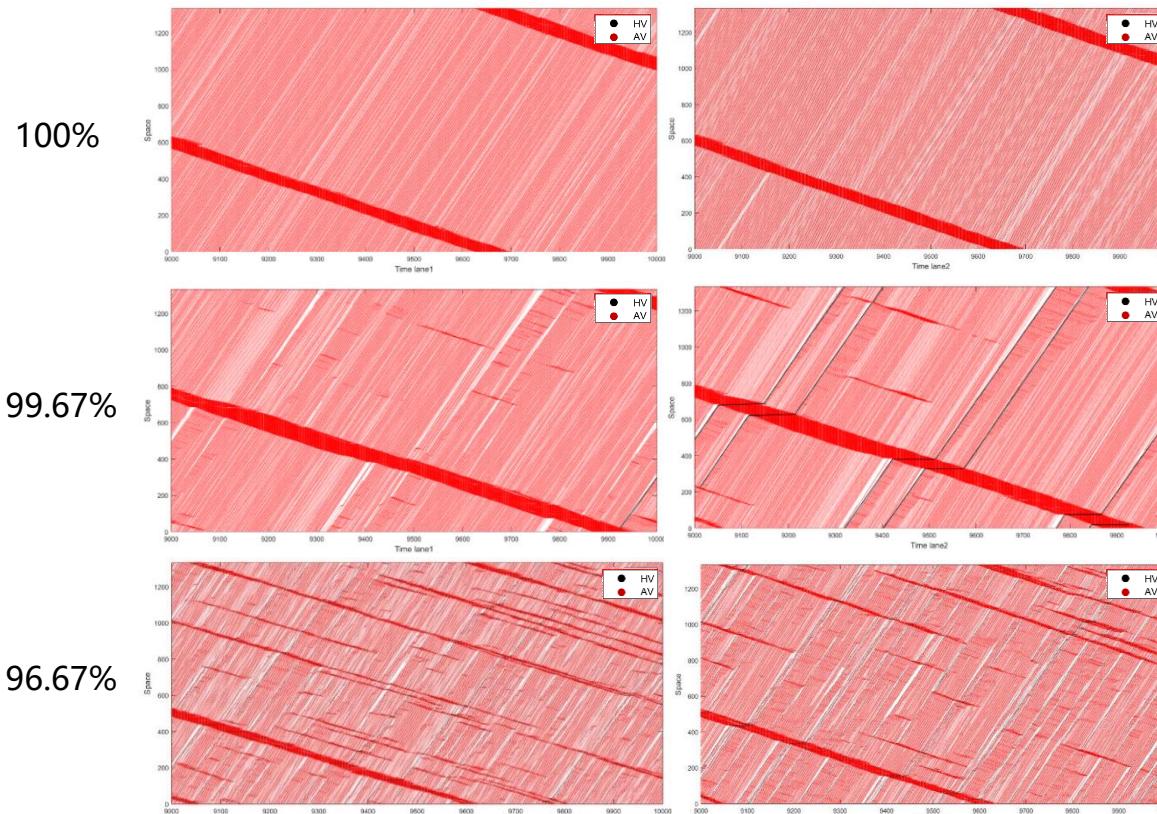
混合交通流特性分析：基本图3-3

当自动驾驶渗透率达到一定高度（比如70%）时，**自动驾驶**对于混合流的效率提升作用不再如之前一般显著，甚至有下降趋势，而此时**人驾车**对混合流的作用占主要部分。因此，**70%**或为最佳自动驾驶渗透率。

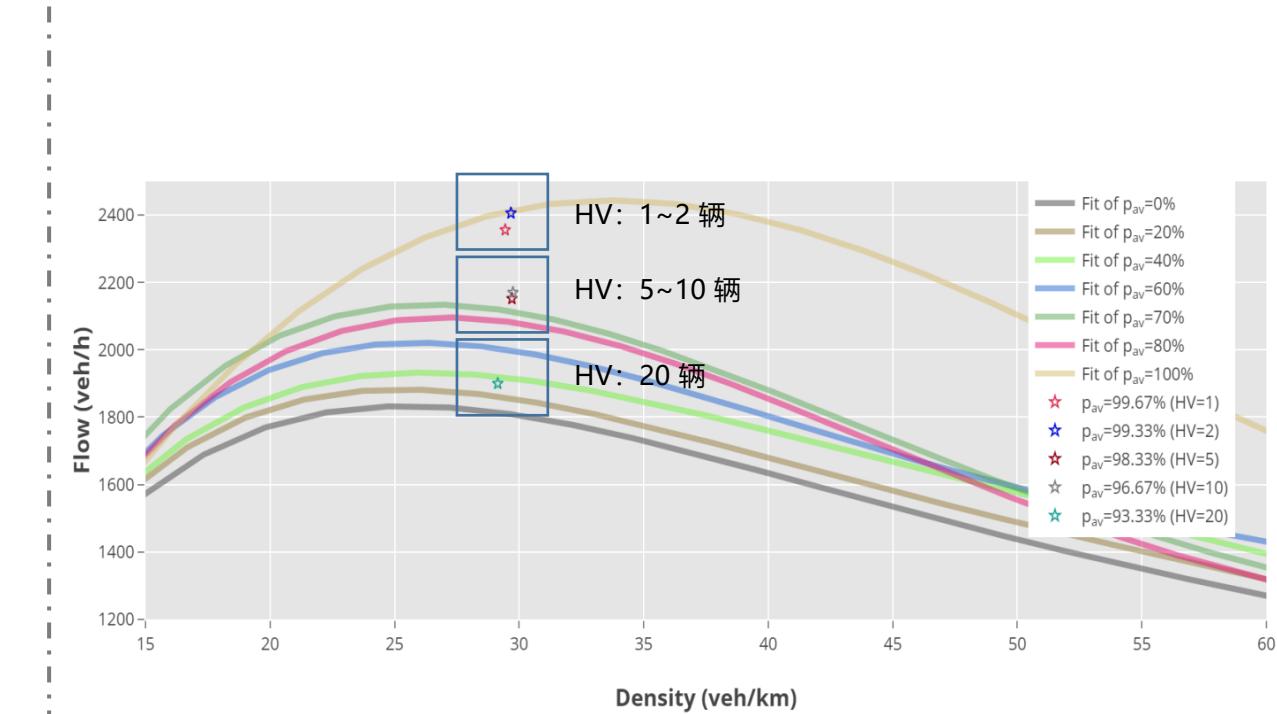


混合交通流特性分析：基本图3-3 (接上页)

少量人驾车在高自动车渗透率下对混合流的关键影响作用



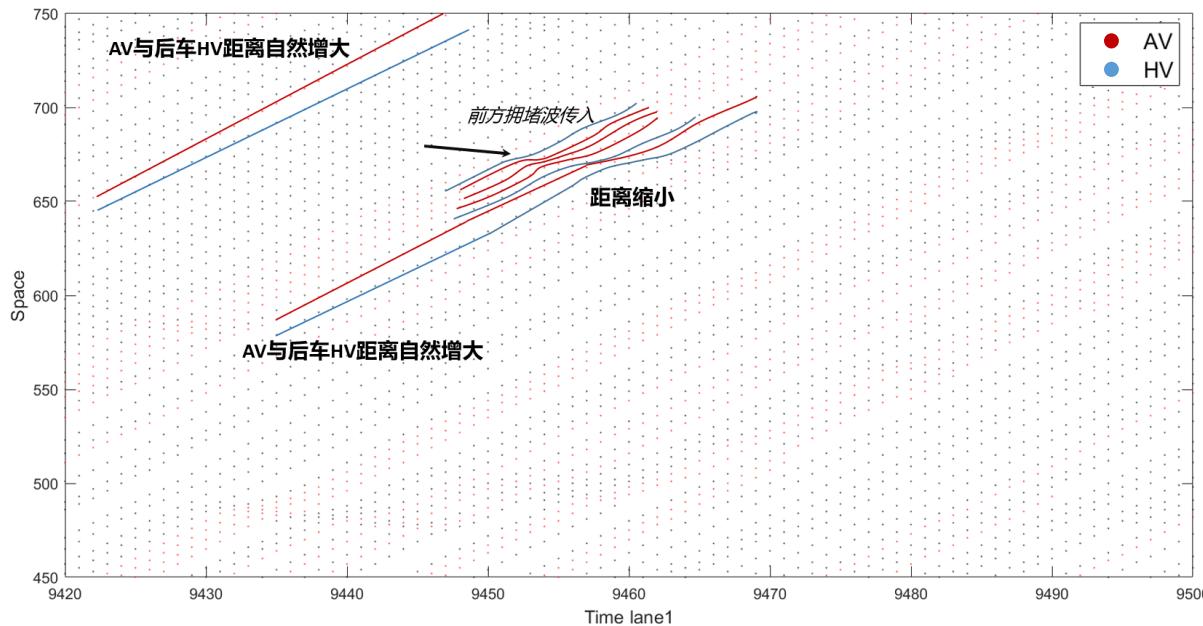
a. 人驾车的混入造成道路空间利用率的下降



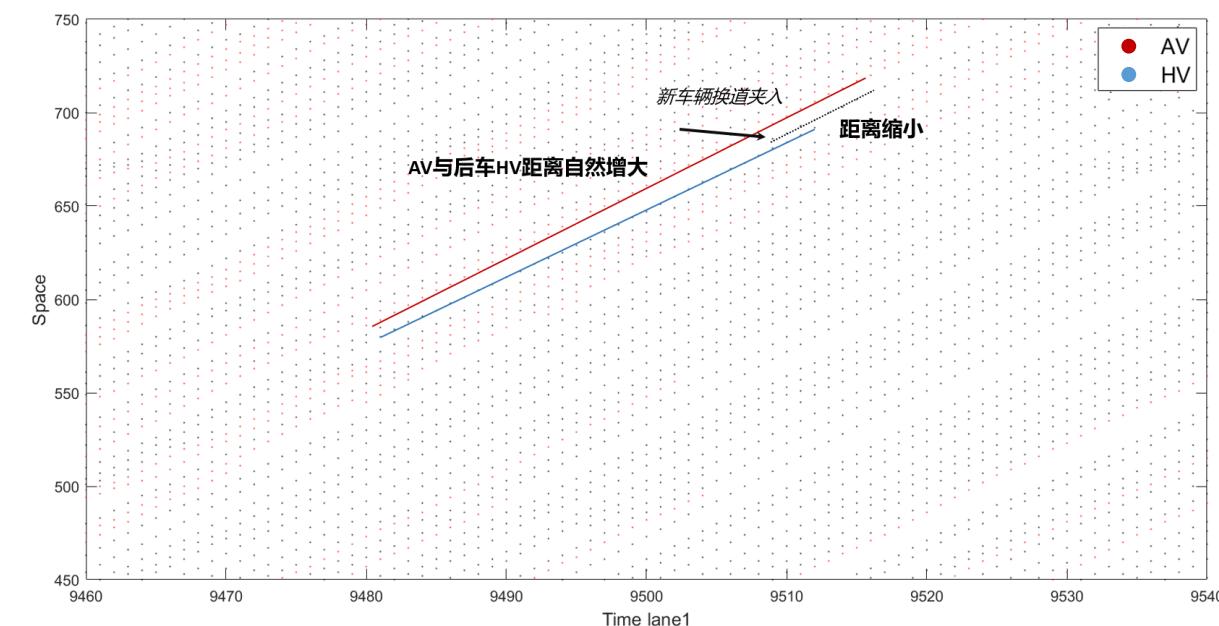
b. 一定范围内，人驾车数量增加，系统流量下降明显

混合交通流特性分析：时空图

自动车相对于人驾车，速度得以提升。除遇以下两种场景：



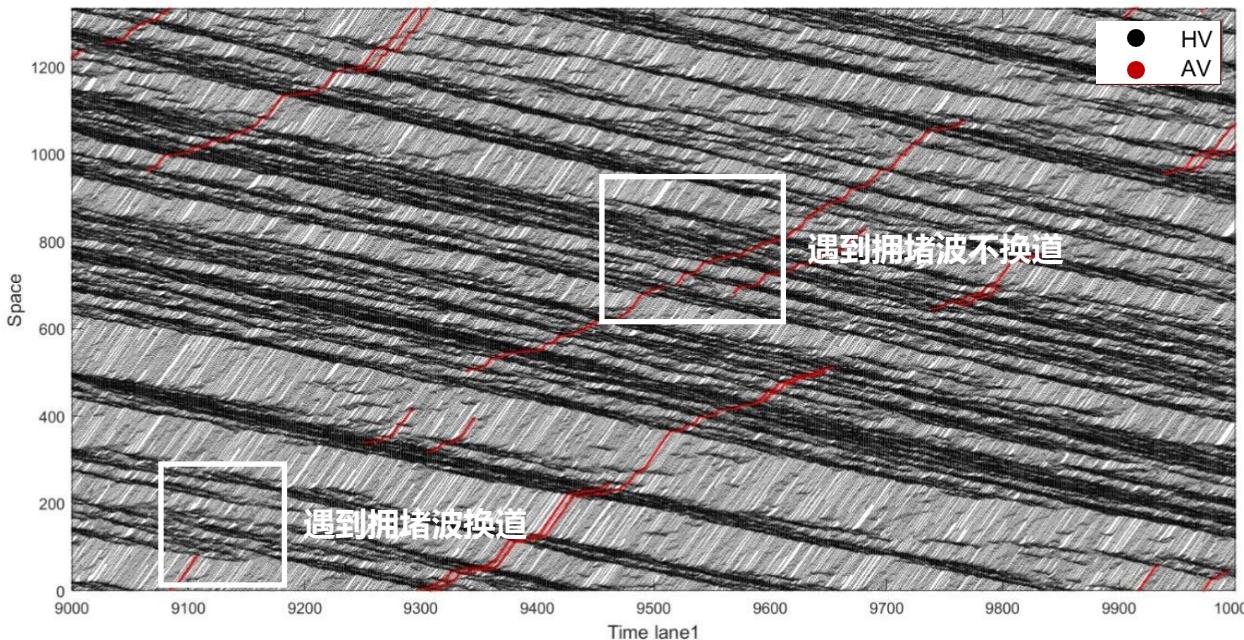
a. 前方拥堵波传入导致的距离缩小



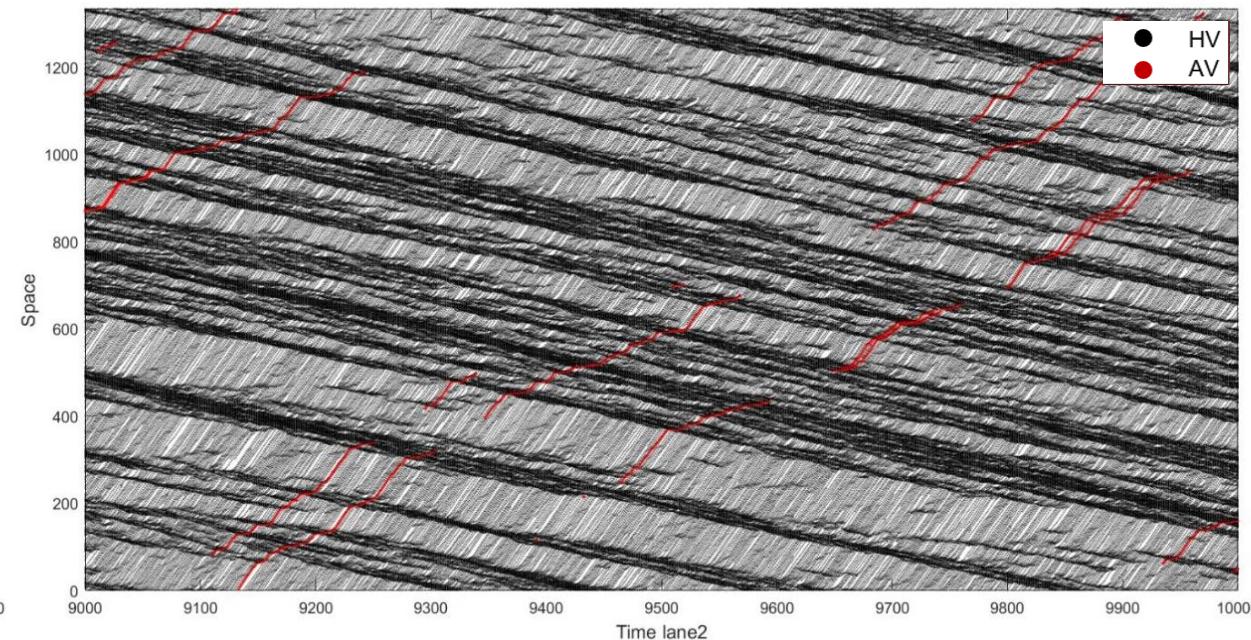
b. 新车辆换道夹入导致的距离缩小

混合交通流特性分析：时空图

当自动驾驶遇到前方传至当前的拥堵波时，会在存在换道空间且换道利于后续行驶时进行换道以提升其自身速度。



车道1

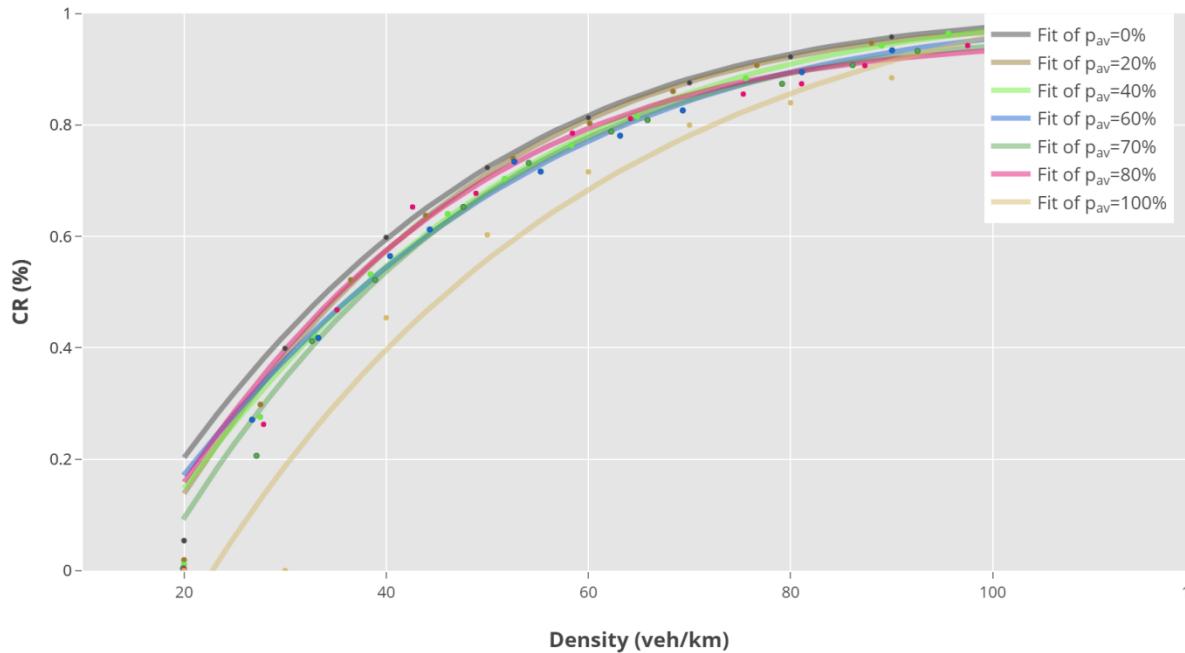


车道2

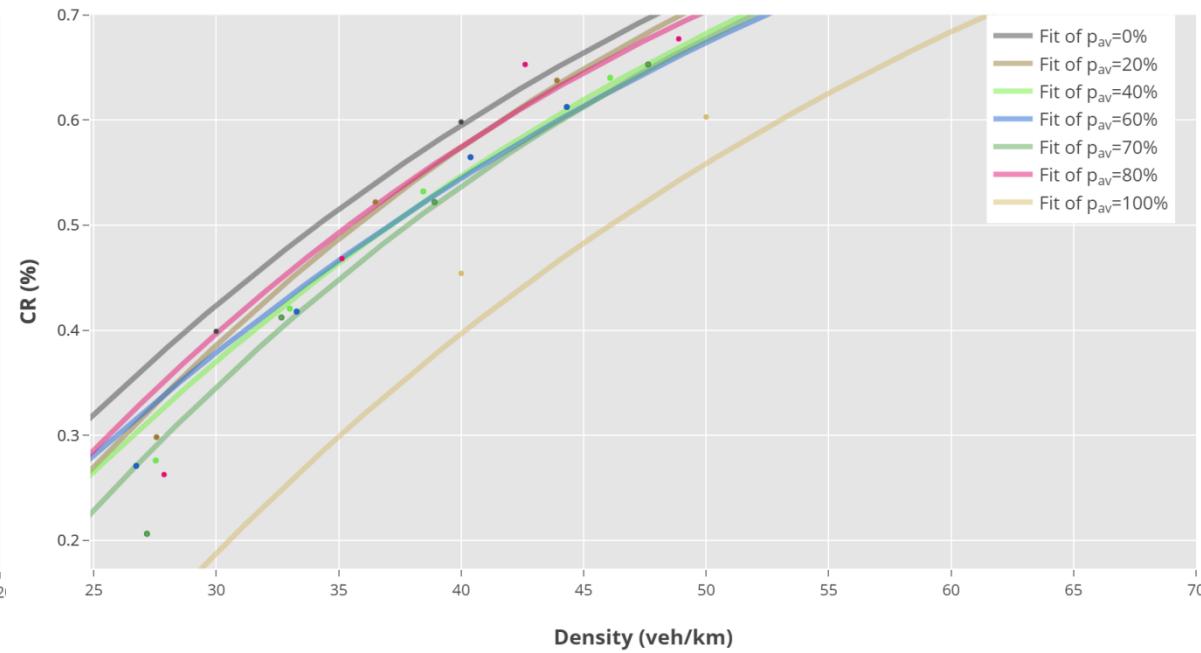
混合交通流特性分析：拥堵比例 CR

$$CR = \frac{n}{\Delta TN}$$

自动车渗透率为0~80%的CR比例曲线接近（相差不超过7%），而纯自动驾驶车辆的CR比例可与纯人驾驶交通流相差最高达到20%。原因在于[人驾车的混入](#)。



不同自动车渗透率下的拥堵比例曲线图（完整）



不同自动车渗透率下的拥堵比例曲线图（局部）

混合交通流特性分析：拥堵比例 CR

$$CR = \frac{n}{\Delta TN}$$

少量人驾车对混合流拥堵程度的较大影响作用。70%自动车渗透率为最优自动车渗透率。

自动车渗透率 (%)	CR (%)
0	39.88
20	39.27
40	36.57
60	33.47
70	28.75
80	30.58
93.33 (20辆HV)	37.13
96.67 (10辆HV)	35.11
98.33 (5辆HV)	31.43
99.33 (2辆HV)	23.80
99.67 (1辆HV)	25.09
100	22.71

相差约10%

相差约15%

拥堵减少

拥堵增加

混合交通流特性分析：平均行驶时间

70%自动车渗透率为最优自动车渗透率。

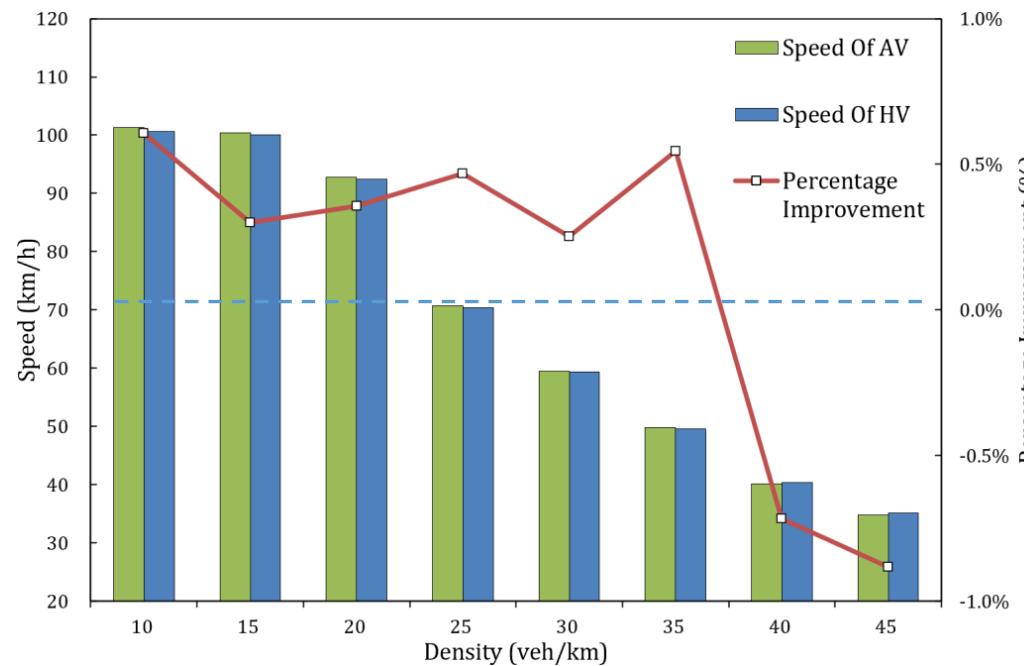
自动车渗透率 (%)	平均行驶时间 (s)
0	172
20	149
40	147
60	146
70	145
80	171
100	139

The diagram illustrates the relationship between car penetration rate and average travel time. It shows a downward-pointing blue arrow from the 70% penetration row to the 145s travel time, accompanied by the text "拥堵增加" (congestion increase). It also shows an upward-pointing blue arrow from the 100% penetration row to the 139s travel time, accompanied by the text "拥堵减少" (congestion decrease).

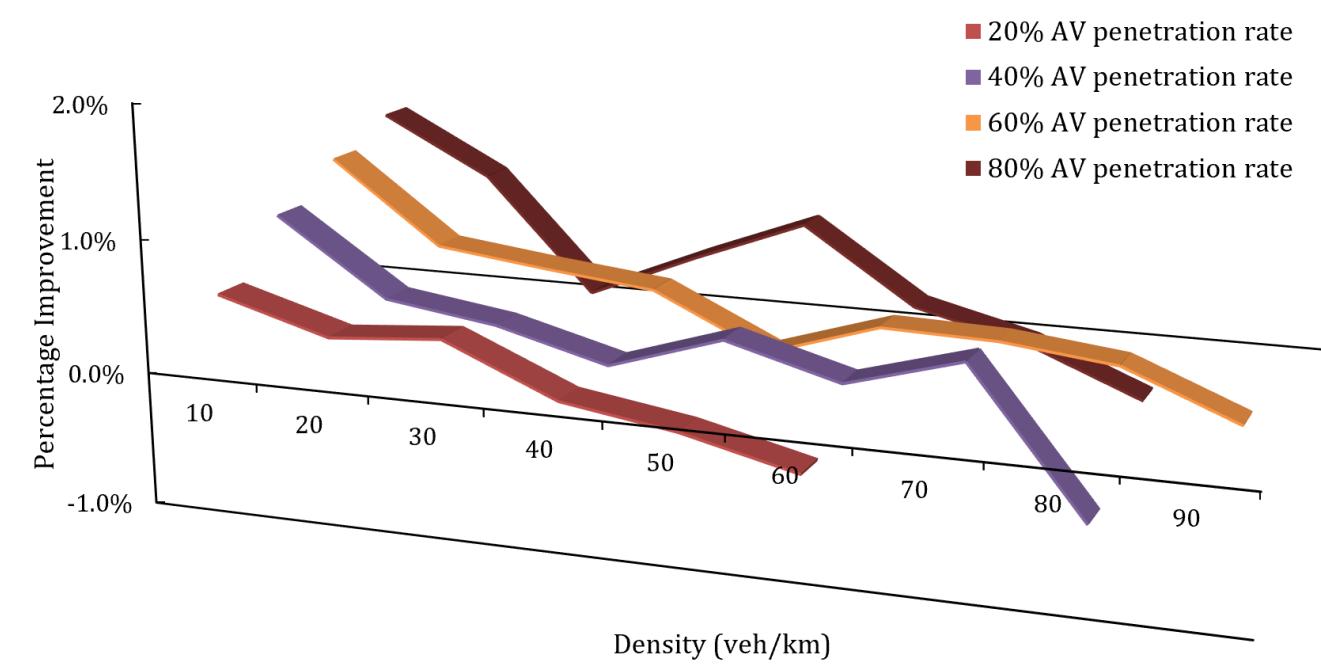
自动驾驶利己策略验证

“无随机慢化”与“有随机慢化”下，自动驾驶利己策略的应用密度范围趋势较为一致：

即自动驾驶的“利己”策略在自由流、临界密度和低拥堵场景下有效



引入随机慢化的自动驾驶与人驾车的平均速度对比



无随机慢化下自动驾驶与人驾车的速度差值百分比变化

自动车利己策略分析

临界密度场景下

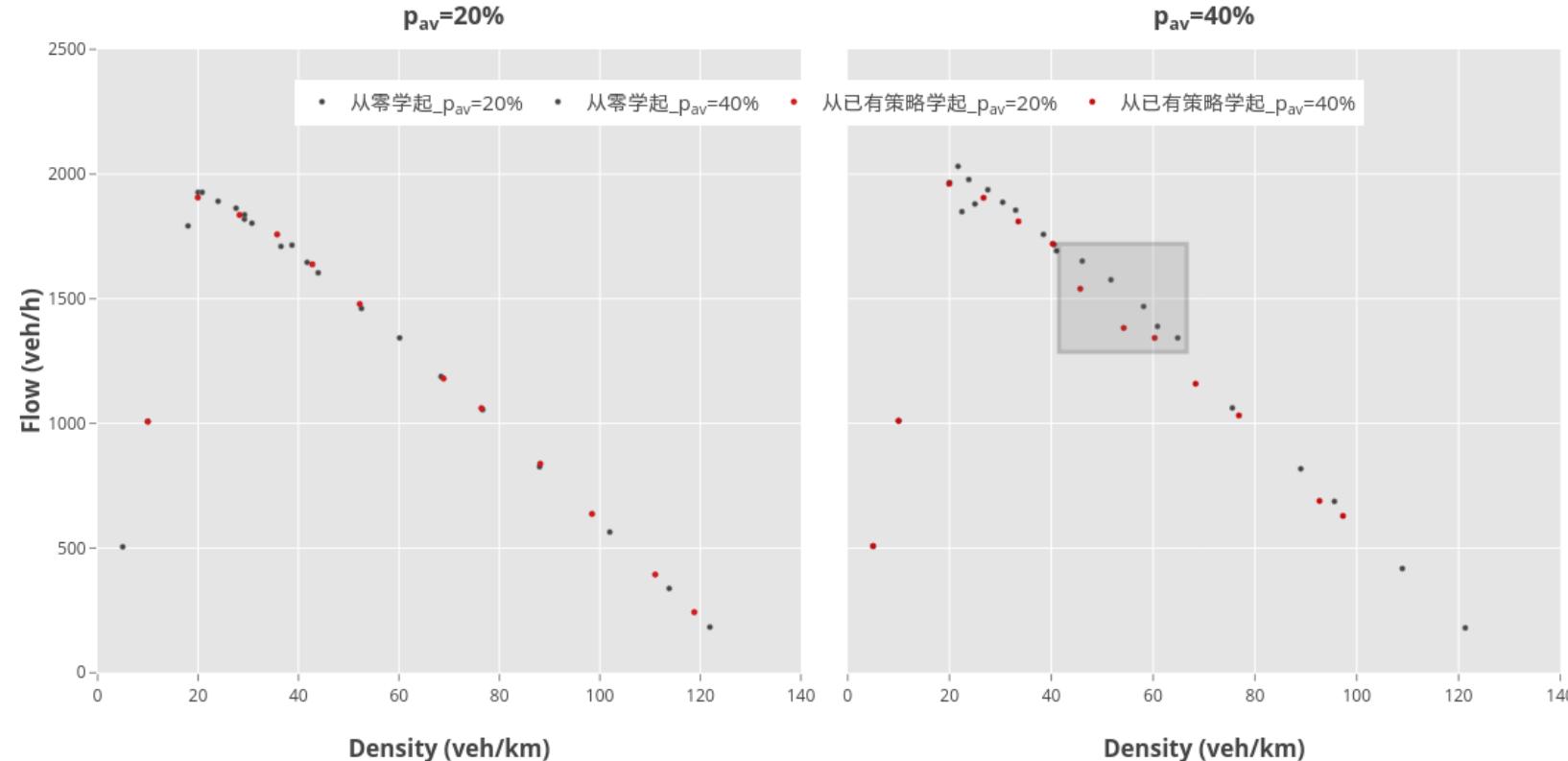
- 强加速**: 与周围车辆有较大相对距离和相对速度时。但自动车渗透率为100%时，自动车较为激进。
- 加速**: 规律性较少。
- 保持**: 自动车渗透率较低时，相对距离与相对速度倾向于集中在1或2；自动车渗透率较高时，相对距离最高频率都在2（即“近距离”至“远距离”范围内），而相对速度最高频率都为1（即靠近）
自动车的加入虽能提高系统运行效率，但“强减速”等动作频率的提升也降低了自动车自身运行的**平稳性**。
- 减速与强减速**: 仅在自动车渗透率为100%时较多发生。
- 换道**: 另一方面，在高自动车渗透率下，人驾车的存在在一定程度上也是自动车运行**不平稳**的原因之一。
另一方面，在高自动车渗透率下，人驾车的存在在一定程度上也是自动车运行**不平稳**的原因之一。

拥堵场景下

- 强加速**: 与临界密度场景下较为一致。但在低自动车渗透率时，中密度下的自动车在更多状态下会采取强加速动作。
- 加速**: 发生在与本车道前车为“近距离”或“中距离”且为“远离”状态时，且随着自动车渗透率的上升，采取“加速”动作的频率有所提高。
- 保持**: 与临界密度场景下较为一致。
- 减速与强减速**: 相较于临界密度下的自动车，拥堵场景下的自动车由于无法始终维持较高速度，因此有更高的频率采取“减速”和“强减速”动作。
- 换道**: 拥堵场景下的“换道”动作比临界密度下的更为激进。

场景适用性分析：同密度，不同自动车渗透率场景

中间密度的训练结果在**低自动车渗透率下**对其他密度完全适用，在**中低自动车渗透率下**对其他密度较适用

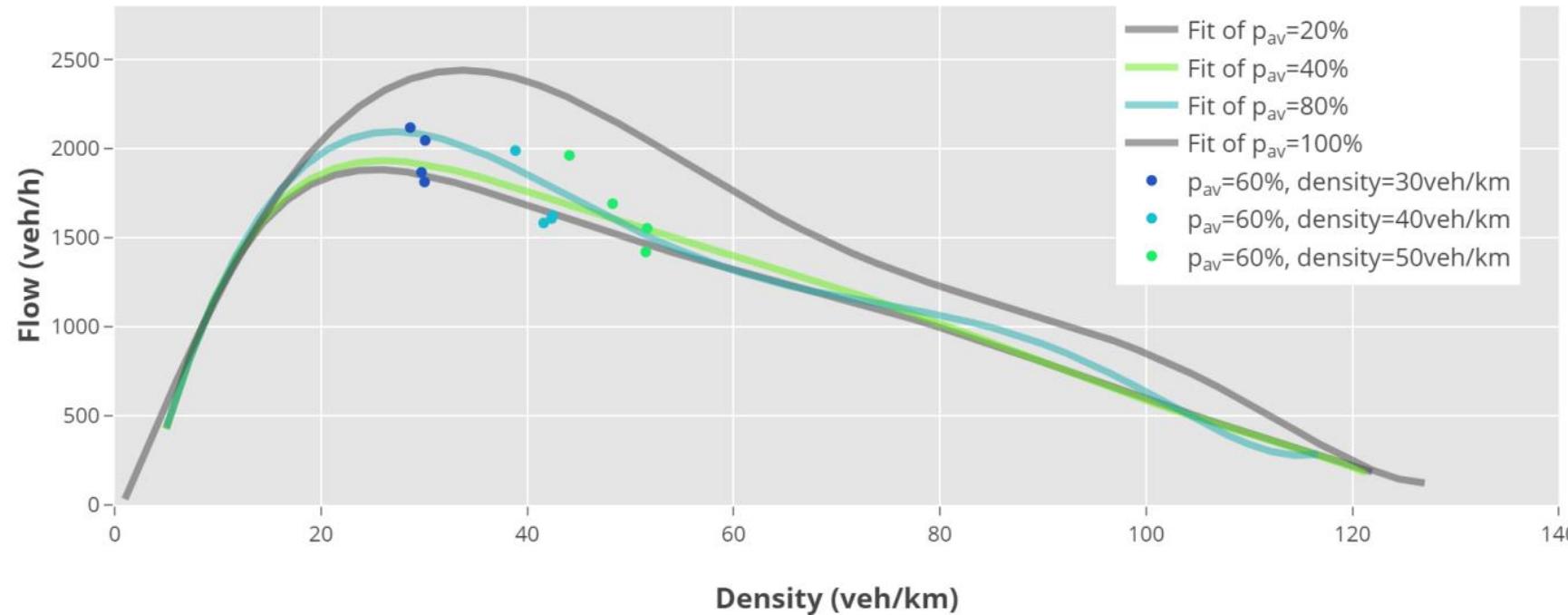


以自动车渗透率为20%和40%为例，将密度为60veh/km的训练结果分别应用至0~125veh/km的场景中

场景适用性分析：不同密度，同自动驾驶渗透率场景

处于中间自动驾驶渗透率的训练结果仅对较低的自动驾驶渗透率适用；

当应用于高自动驾驶渗透率时存在较大误差，特别地，当应用于纯自动驾驶交通流时，误差最大。



以密度为30veh/km、40veh/km和50veh/km为例，将自动驾驶渗透率为60%的训练结果分别应用至自动驾驶渗透率为20%、40%、80%、100%的场景中



4. 论文总结

- 结论
- 主要创新点
- 未来展望

结论

1. 强化学习方法的优越性

- 相对于传统的建模方法，基于强化学习的训练方法更符合汽车行驶的不确定性和智能性。
- 汽车“利己”策略在数值与图形上均得以验证：
 - ✓ 汽车的平均速度在低拥堵场景中大于人驾车。
 - ✓ 汽车通过智能跟车与换道获取长远利益。

2. 最优汽车渗透率

- 在汽车-人驾车混合交通流中，道路通行能力、车流速度、车流稳定性均在一定范围内随着汽车渗透率的增加而提升。
- 汽车渗透率增加到一定程度后，汽车对混合流的正向作用逐渐减弱，而70%被证明是一个可参考的汽车渗透率临界点。

3. 人驾车对混合流的影响

- 在高汽车渗透率（90%~100%）下，少量人驾车对混合流的流量、平均行驶时间、拥堵程度都有较大影响。
- 在混合流中占比0.33%的人驾车对系统整体流量的影响约为其自身比例的13倍。

4. 自动车策略适用性

- 中间密度的训练结果在低汽车渗透率下对其他密度完全适用，在中低汽车渗透率下对其他密度较适用。
- 中间汽车渗透率的训练结果仅对较低的汽车渗透率适用；当应用于高汽车渗透率时存在较大误差，当应用于纯汽车交通流时，误差最大。

主要创新点

1. 强化学习训练技术

- 将强化学习技术作为自动车的策略学习方法，使得学习到的策略不受特定参数的限制，**更加灵活**。
- 该“自我学习”能力符合自动车行驶的智能性与不确定性。

2. “策略适应性”概念

- 提出“策略适应性”概念，并验证了部分场景下的自动车策略可以适用于其他场景，具有一定的**泛化能力**。

3. 元胞自动机与强化学习的结合

- 结合强化学习技术与元胞自动机模型，从多个角度分析自动车对于混合交通流特性的影响，总结可参考的最佳自动车渗透率，提出人驾车在高自动车渗透率下对混合流的关键性影响，为后续**控制自动车与人驾车数量、提升交通系统效率**打下基础。

未来展望

- 适当地修改和丰富元胞尺寸、更新规则，以复现更真实的汽车运行状态；
- 进一步修改道路模型，以求更贴近真实的道路环境；
- 以更小的单位进行仿真测试，填补之前的空缺，以确保更精确的仿真结果；
- 进一步研究自动驾驶车列队对混合流特性的影响。

参考文献

- [1] Bokui Chen, Duo Sun, Jun Zhou, Wengfai Wong, Zhongjun Ding. A future intelligent traffic system with mixed autonomous vehicles and human-driven vehicles[J]. *Information Sciences*, 2020.
- [2] S.D. Pendleton, H. Andersen, X. Du, X. Shen, M. Meghjani, Y.H. Eng, D. Rus, M.H. Ang. Perception, planning, control, and coordination for autonomous vehicles[J]. *Machines*, 2017, 5 (1): 6.
- [3] J Maddox. Improving driving safety through automation, congressional robotics caucus. National Highway Traffic Safety Administration, <https://docplayer.net/10832729-Improving-driving-safety-through-automation.html>, 2012/2020.
- [4] S. Bagloee, M. Tavana, M. Asadi, T. Oliver. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies[J]. *Journal of Modern Transportation*, 2012, 24 (4): 284-303.
- [5] Anderson, K. Nidhi, K. Stanley, P. Sorensen, C. Samaras, O. Oluwatola, Autonomous vehicle technology: A guide for policymakers[EB/OL]. Rand Corporation, https://www.researchgate.net/publication/296697033_Autonomous_Vehicle_Technology_A_Guide_for_Policymakers/citation/download, 2014/2020.
- [6] 苏镜荣,唐翀,程德勇.传统干路与单向二分路通行能力和延误对比[J].*城市交通*,2020,18(02):110-117+134.
- [7] Chen, Z., He, F., Yin, Y., Du, Y.. Optimal design of autonomous vehicle zones in transportation networks[J]. *Transport. Res. Part B: Methodology*, 2017, 99, 44–61.
- [8] Chen, Z., He, F., Zhang, L., Yin, Y.. Optimal deployment of autonomous vehicle lanes with endogenous market penetration[J]. *Transport. Res. Part C: Emerg. Technol.*, 2016, 72, 143–156.
- [9] Levin, M.W., Boyles, S.D.. A cell transmission model for dynamic lane reversal with autonomous vehicles[J]. *Transport. Res. Part C: Emerg. Technol.*, 2016, 68, 126–143.
- [10] Sina Bahrami,Matthew J. Roorda. Optimal traffic management policies for mixed human and automated traffic flows[J]. *Transportation Research Part A*,2020,135.
- [11] T. Litman, Autonomous vehicle implementation predictions[J], Victoria Transport Policy Institute, 2015, No.15-3326.

参考文献

- [12] Danjue Chen, Soyoung Ahn, Madhav Chitturi, David A. Noyce. Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles[J]. *Transportation Research Part B*, 2017, 100.
- [13] Wang, X.R., Jiang, L., Li, Y., Lin, X., Zheng, Wang, F.Y.. Capturing car-following behaviors by deep learning[J]. *IEEE Trans. Intell. Transport.*, 2017, Syst. 99, 1–11.
- [14] 刘赫. 动物行为训练的理论基础 [J]. *中国动物保健*, 2014(2): 23-25.
- [15] 袁耀明. 交通流元胞自动机模型的解析和模拟研究[D]. 中国科学技术大学, 2009.22.
- [16] M. Cremer, J. Ludwig, A fast simulation model for traffic flow on the basis of boolean operations[J]. *Math. Comput. Simul.*, 1986, 28 (4): 297– 303.
- [17] K. Nagel, M. Schreckenberg. A cellular automaton model for freeway traffic[J]. *Phys. I France*, 1992, 2: 2221–2229.
- [18] O. Biham, A.A. Middleton, D. Levine. Self-organization and a dynamical transition in traffic-flow models[J]. *Phys. Rev. A*, 1992, 46 (10): R6124–R6217.
- [19] M. Takayasu, H. Takayasu. 1/f noise in a traffic model[J]. *Fractals*, 1993, 1 (4): 860–866.
- [20] K. Nagel, M. Paczuski. Emergent traffic jams[J]. *Phys. Rev. E*, 1995, 51 (4): 2909–2918.
- [21] Knospe W, Schadschneider A, Schreckenberg M, et al. Towards a realistic microscopic description of highway traffic[J]. *Journal of Physics A General Physics*, 2000, 33(48): L477.
- [22] Li X, Wu Q, Jiang R. Cellular automaton model considering the velocity effect of a car on the successive car[J]. 2001, 64(6 Pt 2):066128.
- [23] T. Nagatani. Self-organization and phase transition in traffic-flow model of a two-lane roadway[J]. *Phys. A: Math. Gen.*, 1993, 26: 781.
- [24] M. Rickert, K. Nagel, M. Schreckenberg, A. Latour. Two lane traffic simulations using cellular automata[J]. *Phys. A*, 1996, 231: 534.
- [25] P. Wagner, K. Nagel, D.E. Wolf. Realistic multi-lane traffic rules for cellular automata[J]. *Phys. A*, 1997, 234: 687–698.
- [26] Arnab Bose, Petros Ioannou. Mixed manual/semi-automated traffic: a macroscopic analysis[J]. 2002, 11(6):439-462.
- [27] Jincai Chang, Zhuo Wang, Tong Xiao, et al. Modeling and simulations on automated vehicles to alleviate traffic congestion[J]. 2017, 3(2):112-125.
- [28] Danjue Chen, Soyoung Ahn, Madhav Chitturi, et al. Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles[J]. 2017, 100:196-221.
- [29] Sina Bahrami, Matthew J. Roorda. Optimal traffic management policies for mixed human and automated traffic flows[J]. 2020, 135:130-143.

参考文献

- [30] Jiazu Zhou, Feng Zhu. Modeling the fundamental diagram of mixed human-driven and connected automated vehicles[J]. 2020, 115.
- [31] Yangzexi Liu, Jingqiu Guo, John Taplin, et al. Characteristic Analysis of Mixed Traffic Flow of Regular and Autonomous Vehicles Using Cellular Automata[J]. 2017.
- [32] Simonelli, F., Bifulco, G., De Martinis, V., Punzo, V.. Human-like adaptive cruise control systems through a learning machine approach[J]. Appl. Soft Comput.,2009,240–249.
- [33] Meixin Zhu, Xuesong Wang, Yinhai Wang. Human-like autonomous car-following model with deep reinforcement learning[J]. Transportation Research Part C,2018,97.
- [34] 山岩. 自动车拟人化换道决策和换道轨迹研究[D].长安大学, 2019.
- [35] LITTMAN M L. Reinforcement learning improves behaviour from evaluative feedback[J]. Nature, 2015, 521(7553): 445-451.
- [36] J. Guo, S. Cheng and Y. Liu. Merging and Diverging Impact on Mixed Traffic of Regular and Autonomous Vehicles[J]. IEEE Transactions on Intelligent Transportation Systems,2020.
- [37] Changxi You, Jianbo Lu, Dimitar Filev, Panagiotis Tsotras. Advanced planning for autonomous vehicles using reinforcement learning and deep inverse reinforcement learning[J]. Robotics and Autonomous Systems,2019.
- [38] CHOWDHURY D, WOLF D E, SCHRECHENBERG M. Particle hopping models for two-lane traffic with two kinds of vehicles: effects of lane-changing rules[J]. Phys. A: Statistical Mechanics and its Applications, 1997, 235(3/4): 417-439.
- [39] Sven Maerivoet, Bart De Moor. Cellular automata models of road traffic[J]. Physics Reports, 2005, 419: 1-64.
- [40] K. Nagel, H.J. Herrmann, Deterministic models for traffic jams[J]. Physics. A, 1993, (199): 254.
- [41] Yangzexi L., Jingqiu G., John T., et al. Characteristic Analysis of Mixed Traffic Flow of Regular and Autonomous Vehicles Using Cellular Automata[J]. Journal of Advanced Transportation, 2017, 2017:1-10.
- [42] 郭静秋,方守恩,曲小波,王亦兵,刘洋泽西.基于强化协作博弈方法的双车道混合交通流特性[J].同济大学学报(自然科学版),2019,47(07):976-983.
- [43] Mohsen Kamrani, Aravinda Ramakrishnan Srinivasan,Subhadeep Chakraborty,Asad J. Khattak. Applying Markov decision process to understand driving decisions using basic safety messages data[J]. Transportation Research Part C,2020,115.

参考文献

- [44] R. Bellman. A Markovian decision process[J]. *Math. Mech.*, 1957, 679–684.
- [45] R.S. Sutton, A.G. Barto. Reinforcement Learning: An Introduction[M], vol. 1, no. 1, MIT Press, Cambridge, 1998.
- [46] P. Brémaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*[J], Springer Science & Business Media, 2013, vol. 31.
- [47] 赵婷婷,孔乐,韩雅杰,任德华,陈亚瑞.模型化强化学习研究综述[J/OL].*计算机科学与探索*:1-11.<http://kns.cnki.net/kcms/detail/11.5602.tp.20200331.1819.002.html>, 2020-05-12.
- [48] 赵冬斌,邵坤,朱圆恒,李栋,陈亚冉,王海涛,刘德荣,周彤,王成红.深度强化学习综述:兼论计算机围棋的发展[J].*控制理论与应用*,2016,33(06):701-717.
- [49] Wang S.-C. Artificial neural network[M]. *Interdisciplinary Computing in Java Programming*, Springer, 2003, pp. 81-100.
- [50] Bernardo J.M., Smith A.F. *Bayesian Theory*[M]. John Wiley & Sons, Canada, 2001.
- [51] Shi J.Q., Choi T. *Gaussian Process Regression Analysis for Functional Data*[M]. CRC Press, 2011.
- [52] Cristianini N., Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*[M].Cambridge University Press, 2000.
- [53] C.J.C.H. Watkins. *Learning from Delayed Rewards* (Ph.D. dissertation) [D]. King's College, Cambridge, 1989.



北京工业大学本科毕业论文结题（申优）答辩

自动驾驶车辆利己行为建模与影响分析

恳请各位评委老师批评指正

答辩人：李悦（交通设备与控制工程 16030120）