7CS075/UZ1: Research Methods

Intelligent Agents in Autonomous Driving

## **1.Introduction**

Russell and Norvig, (2016) defined Artificial Intelligence (AI) in four main categories:

1. Thinking humanly
2. **A**cting humanly
3. Thinking rationally
4. Acting rationally

Thinking humanly involves cognitive modeling to emulate human thought processes, while acting humanly is evaluated by the Turing Test, which assesses whether machines can mimic human behavior (Russell and Norvig, 2016;Hodges, 2012). Thinking rationally focuses on logic-based systems that follow formal rules, and acting rationally involves optimising actions to achieve specific goals (Russell and Norvig, 2016). AI aims to perform tasks requiring human intelligence, such as learning, pattern recognition, and decision-making.

AI employs techniques like machine learning, neural networks, and natural language processing to create adaptive systems. Machine learning includes supervised, unsupervised, and reinforcement learning (RL), which involves agents learning from environmental interactions. Neural networks assist with pattern recognition and decision-making, while natural language processing enables interaction with human language.

In autonomous vehicles (AVs), intelligent agents are crucial for perception, decision-making, and autonomous action. These include reflex agents, model-based agents, goal-based agents, utility-based agents, and RL agents (Russell and Norvig, 2016). This report evaluates how RL algorithms and model-based agents address the challenge of autonomous driving. Understanding and integrating these agents are essential for enhancing AV safety, efficiency, and reliability.

## **2.Aim**

Assess the application of reinforcement learning versus model-based agents in solving autonomous vehicle challenges.

### 2.1 Objectives

1. Explore the fundamentals of RL and model-based agent architectures and provide a comprehensive overview.
2. Review existing research and applications of these agents in autonomous vehicle technology.
3. Assess the strengths and limitations of RL and model-based agents in autonomous driving scenarios and explore potential hybrid approaches for enhanced performance in autonomous vehicles.
4. Draw conclusions on the optimal strategies for integrating intelligent agents in autonomous vehicle systems to ensure safety, efficiency, and reliability.

Achieving these objectives is crucial for overcoming challenges in the AV industry. A solid understanding of RL and model-based agent architectures will propel innovative research. Evaluating their strengths and limitations ensures a balanced assessment of both architectures. Exploring hybrid methods can combine the benefits of RL and model-based agents, enhancing AV performance. The performance of AVs can be measured using metrics such as safety (collision rates), efficiency (travel time and energy consumption), reliability (system uptime and fault tolerance), and user satisfaction (comfort and acceptance).

## **3. Literature Review**

### 3.1 Reinforcement Learning

RL is a **model-free** approach where agents learn optimal behaviors through interactions with the environment, receiving rewards or penalties based on their actions (Sutton and Barto, 2018). Key RL algorithms such as Q-Learning, Deep Q-Networks (DQN), and Proximal Policy Optimisation (PPO) have demonstrated effectiveness in AV systems, particularly in complex driving scenarios. Q-Learning, a foundational RL algorithm, updates Q-values—estimates of the expected utility of actions—using the Bellman equation (Sutton and Barto, 2018). It has been applied to simpler driving tasks like lane-keeping and obstacle avoidance. However, Q-Learning struggles with high-dimensional state spaces and complex scenarios (Van Hasselt, Guez and Silver, 2016).

To address these limitations, DQN incorporates deep learning techniques to approximate Q-values with a neural network, enabling the handling of complex inputs such as images. DQN has shown effectiveness in tasks requiring visual perception and complex decision-making, such as enhancing lane-keeping accuracy in AV systems (Khanum, Lee and Yang, 2022). PPO improves earlier RL methods by enhancing stability and reliability in policy updates through a clipping mechanism and a surrogate objective function (Kendall et al., 2019). It is well-suited for continuous action spaces and dynamic environments, as demonstrated in end-to-end driving systems where PPO aids in managing complex driving scenarios with precision (Kendall et al., 2019). Despite these strengths, RL methods face challenges including sample inefficiency, high computational costs, and issues with sparse or delayed rewards, which impact training stability and real-time application feasibility.

RL has been employed to solve various practical problems in autonomous driving, each with unique challenges and requirements. One common scenario is lane-keeping, where the AV must maintain its position within a lane while accounting for lane curvature, road boundaries, and other vehicles. Q-Learning can handle this task in simpler environments by learning policies to stay centered in the lane, but its effectiveness diminishes in more complex scenarios with high-dimensional inputs.

A more challenging scenario involves navigating intersections, which requires the AV to handle dynamic elements like traffic lights, pedestrian crossings, and other vehicles entering or exiting the intersection. Here, DQN excels by processing visual inputs from cameras and lidar sensors to make split-second decisions. It can learn to recognise and react to traffic light changes, yielding to pedestrians, and avoiding collisions with other vehicles, while optimising for safety and efficiency (Khanum, Lee and Yang, 2022)

Emergency scenarios, such as obstacle avoidance, present additional challenges where RL can be highly beneficial. If an unexpected obstacle appears on the road, the AV must quickly decide whether to brake, steer around the obstacle, or take another evasive action. RL algorithms like DQN and PPO can learn from simulated environments to execute these maneuvers with precision, reducing the risk of accidents and improving overall safety.

RL can optimise driving strategies in dynamic urban environments by handling complex tasks like route planning, speed adjustment, and interactions with unpredictable elements such as pedestrians and other drivers. RL adapts to new situations, enabling AVs to navigate efficiently and safely, thus improving their reliability and performance in real-world scenarios. Despite challenges like sample inefficiency and high computational costs, advances in RL algorithms and computing power are enhancing the feasibility and effectiveness of RL in AV systems.

In the context of the Unity project, an autonomous agent is designed to navigate and respond dynamically to environmental stimuli. By employing RL, the agent continuously improves its decision-making process based on rewards and penalties received for its actions. This iterative learning process allows the agent to adapt to complex and changing scenarios, ultimately enhancing its performance in real-world applications.

### 3.2 Model Based Agents

**Model-based agents** utilise an internal model of the environment to simulate outcomes and make decisions, focusing on predictive planning essential for autonomous driving. **Monte Carlo tree search** (MCTS) is a heuristic search algorithm that builds a search tree incrementally through simulations of random outcomes to evaluate potential decisions (Mo, Pei and Wu, 2021). MCTS is effective in large state spaces where exact solutions are impractical, aiding in decision-making tasks like path planning and maneuver selection in AV systems (Mo, Pei and Wu, 2021). **Dynamic Programme** (DP) addresses problems by breaking them into simpler subproblems and storing results to avoid redundant calculations (Mo, Pei and Wu, 2021). Techniques such as value iteration and policy iteration are used to update state values and improve policies iteratively. DP is effective for managing state spaces and has been applied to route planning and resource allocation in AV systems (Russell and Norvig, 2016). By incorporating a model-based approach, the Unity agent can simulate realistic and intelligent decision-making processes, enhancing its capability to navigate complex environments and respond effectively to unforeseen events.

In scenarios requiring long-term planning such as merging onto highways or navigating through complex intersections, model-based methods can effectively anticipate and simulate potential traffic situations, thereby helping the AV make informed decisions (Russell and Norvig, 2016). However, their effectiveness is limited by the accuracy of the internal model and the high computational complexity of simulations, which may restrict real-time application feasibility. Comparatively, model-based approaches such as **Model Predictive Control** (MPC) can handle real-time constraints better than traditional DP but still face computational challenges (Mo, Pei and Wu, 2021).

#### **4.Analysis of Findings**

The choice between RL and model-based approaches depends on the specific needs of the driving environment. RL methods are flexible and handle complex tasks but require significant computational resources and training time. Model-based agents offer precise predictions but can be limited by model accuracy and computational demands. Combining RL with model-based methods can harness the strengths of both, improving overall performance and functionality in autonomous driving systems but also presents challenges related to model accuracy and computational complexity (Ha and Jürgen Schmidhuber, 2018; Chua et al., 2018; Kendall et al., 2019). A hybrid system might use model-based planning for strategic decisions and RL for tactical maneuvers, optimising both long-term and short-term performance.

#### **5.Conclusion**

RL provides adaptability and continuous improvement, making it ideal for dynamic driving environments. In contrast, model-based agents offer high accuracy in prediction and planning, crucial for making informed decisions. Integrating RL's learning capabilities with the predictive precision of model-based agents represents a promising strategy for advancing autonomous driving technology, ensuring improved performance and safety.

There are several ethical implications that should be carefully considered before the application of AI within AVs. Decision-making in life-critical situations, such as unavoidable accidents, poses significant ethical dilemmas. How should an AV prioritise the safety of its passengers versus pedestrians? Privacy concerns also arise from the vast amount of data AVs collected and processed. Ensuring data security and protecting user privacy are crucial for gaining public trust in AV technology.

# Reference list

1. Chua, K., Calandra, R., McAllister, R. and Levine, S. (2018). Deep Reinforcement Learning in a Handful of Trials Using Probabilistic Dynamics Models. *ArXiv (Cornell University)*, 31. doi: https://doi.org/10.48550/arxiv.1805.12114.
2. Ha, D. and Jürgen Schmidhuber (2018). World Models. *ArXiv (Cornell University)*. doi: https://doi.org/10.48550/arxiv.1803.10122.
3. Hodges, A. (2012). *Alan Turing: The Enigma*. Princeton, N.J.: Princeton University Press, Princeton, New Jersey.
4. Kendall, A., Hawke, J., Janz, D., Mazur, P., Reda, D., Allen, J.-M., Lam, V.-D., Bewley, A. and Shah, A. (2019). *Learning to Drive in a Day*. [online] IEEE Xplore. doi: https://doi.org/10.1109/ICRA.2019.8793742.
5. Khanum, A., Lee, C.-Y. and Yang, C.-S. (2022). Deep-Learning-Based Network for Lane following in Autonomous Vehicles. *Electronics*, 11(19), p.3084. doi: https://doi.org/10.3390/electronics11193084.
6. Mo, S., Pei, X. and Wu, C. (2021). Safe Reinforcement Learning for Autonomous Vehicle Using Monte Carlo Tree Search. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), pp.1–8. doi: https://doi.org/10.1109/tits.2021.3061627.
7. Russell, S. and Norvig, P. (2016). *Artificial Intelligence: A Modern Approach, Global Edition*. 3rd ed. Boston, Mass: Pearson.
8. Sutton, R.S. and Barto, A. (2018). *Reinforcement Learning: an Introduction*. Cambridge, Ma; Lodon: The Mit Press.
9. Van Hasselt, H., Guez, A. and Silver, D. (2016). Deep Reinforcement Learning with Double Q-Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1). doi: https://doi.org/10.1609/aaai.v30i1.10295.