ICRA LONDON · 2023

29 May – 2 June 2023 | ExCeL London

EMBRACING THE FUTURE: MAKING ROBOTS FOR HUMANS

ICRA Automation Cluster Forum

Artificial Intelligence for Autonomous Vehicles Control

Maria Pia Fanti,

Lab of Control and Automation, Department of Electrical and Information Engineering, Polytechnic University of Bari, Italy

IEEE RAS Technical committee of Automation in Logistics

Politecnico di Bari

London, May 30, 2023

LCA

DEI DIPARTIMENTO DI INGEGNERIA ELETTRICA E DELL'INFORMAZIONE

OUTLINE

Machine learning (ML) and deep reinforcement learning (DRL) techniques applied in autonomous vehicle control:

- **traffic management** (efficiently handling by intelligent traffic lights road traffic in road intersections where priority issues are important)
- **traffic management** (crossroads without traffic lights involving connected automated vehicles, priority vehicles and regular vehicles)
- **automotive** (autonomous braking systems based on an intelligent agent)

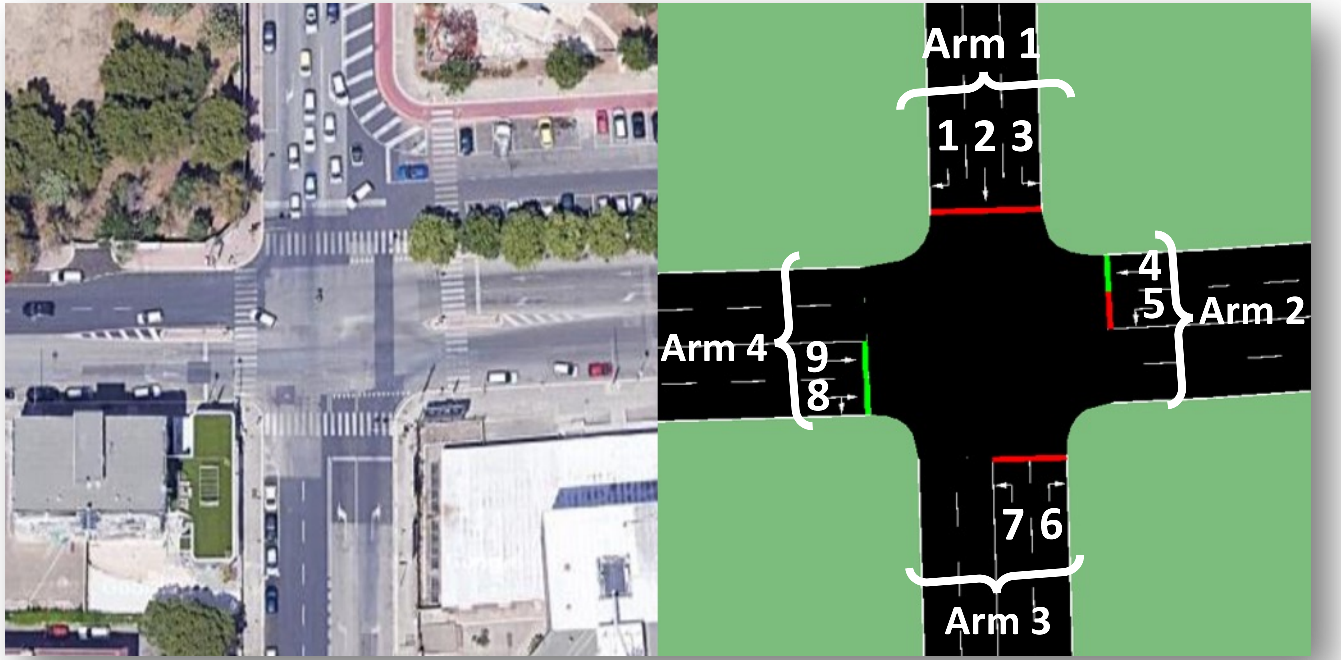
OUTLINE

Machine learning (ML) and deep reinforcement learning (DRL) techniques applied in autonomous vehicle control:

- **traffic management** (efficiently handling by intelligent traffic lights road traffic in road intersections where priority issues are important)
- **traffic management** (crossroads without traffic lights involvingconnected automated vehicles, priority vehicles and regular vehicles)
- **automotive** (autonomous braking systems based on an intelligent agent)

Problem Statement

A control strategy to select the traffic light phases with the main objective of managing three classes of priority for vehicles.



Three priority classes:

P_1: low-priority vehicles (Cars)

P_2 : intermediate-priority vehicles (Buses)

P_3 : high-priority vehicles (ambulances and emergency vehicles)

SUMO
SIMULATION OF URBAN MOBILITY

An intelligent agent is trained to manage the traffic light system of the road intersection in order to improve the performance in terms of waiting time.

Deep Reinforcement Learning Framework

Deep Reinforcement Learning deterministic model: **the triple (S,A,R)**
S : **State Space,** the set of the possible states of the environment
$s_t \in S$ is the state at the time-step t

A: **Action Space**, the set of the possible agent actions in the environment
$a_t \in A$ is the action at time step t

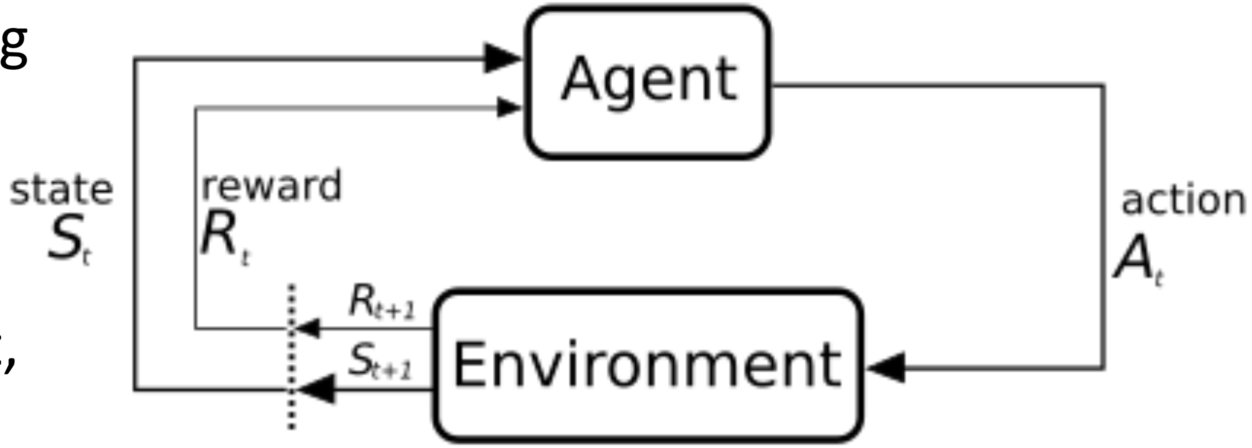
R : **Reward Space,** a set of possible rewards that an agent can gain as consequence of its action on the environment
$r(s_t, a_t) \in R$ is the reward obtained with the action a in state s in time-step t.

Deep Reinforcement Learning Framework

Performing action a_t at time t, the agent makes the environment to transit from s_t to s_{t+1}, gaining a reward r_t as a measure of the effects of its action chosen on the environment.

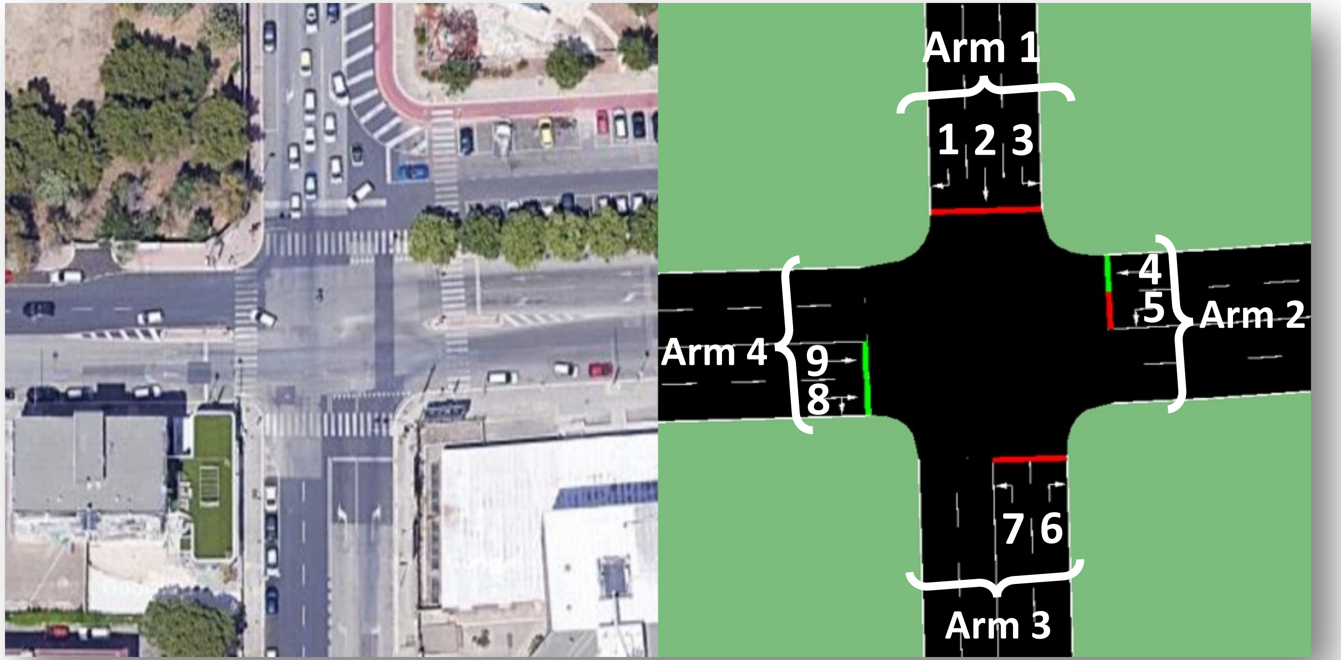
The policy p_i is the way the agent decides to act, thus choosing its actions a_t.

The choice of the action is performed on the basis of a Q- Function Q(s, a).



Case study

The control strategy is applied to a case study modelling a road intersection located in the city center of Bari (Italy).



Vehicle length: 3,5 m
Vehicle speed: 30 km/h

State space

Intersection model:
a set of 9 lanes, each lane is divided into 40 cells, each cell has a length of 3,5m.

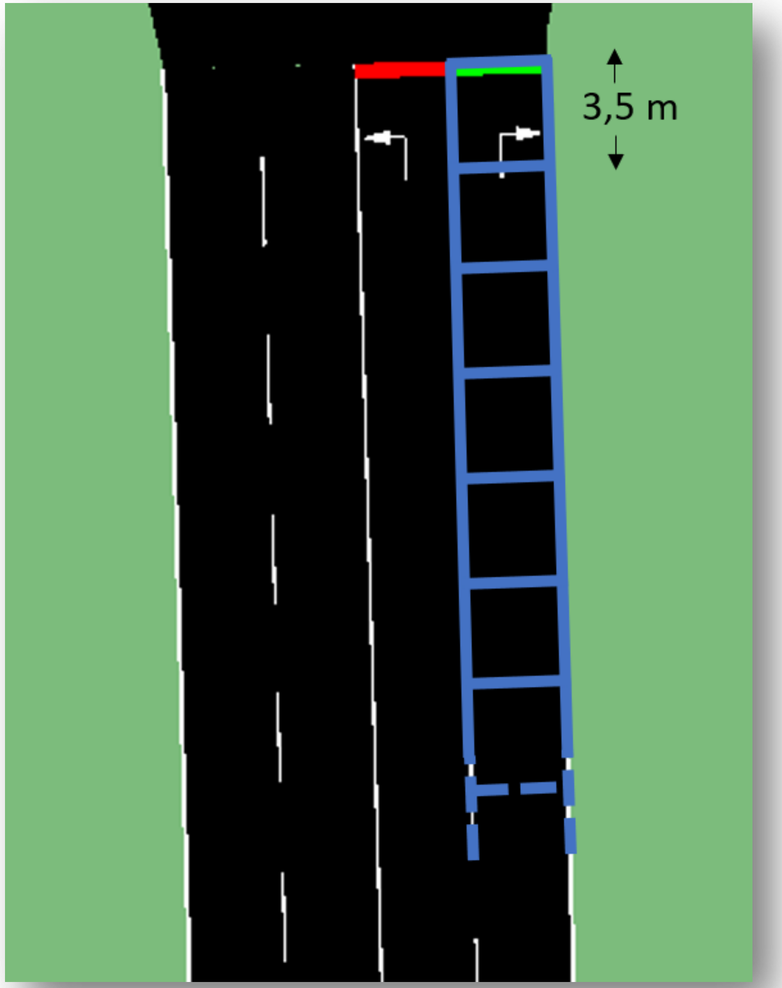
State of each lane ***i***:

$$\boldsymbol{s\_lane}^{(i)} = \left[x_k^{(i)}\right], k = 1, \dots, 40;$$

- $x_k^{(i)} = 0$ if empty;
- $x_k^{(i)} = 1$ if p_1 vehicle;
- $x_k^{(i)} = 10$ if p_2 vehicle;
- $x_k^{(i)} = 100$ if p_3 vehicle.

State matrix $\boldsymbol{s_t}$:

$$\boldsymbol{s_t} = \begin{bmatrix} \boldsymbol{s\_lane}^{(1)} \\ \vdots \\ \boldsymbol{s\_lane}^{(i)} \\ \vdots \\ \boldsymbol{s\_lane}^{(9)} \end{bmatrix}$$



Action space

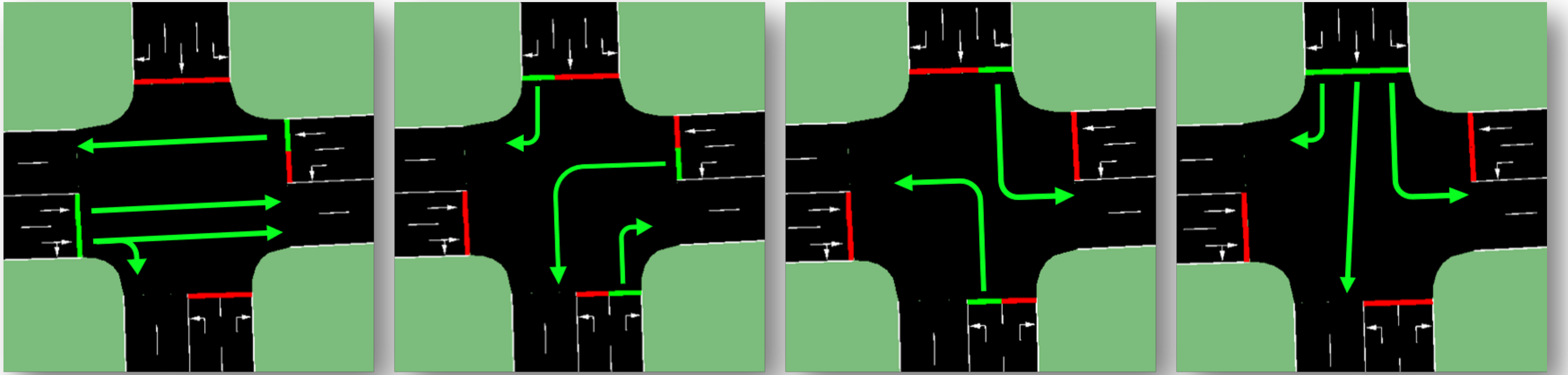
At each step, the agent can choose one of the following actions:

1 (a): green phase for lanes 4, 8 and 9, and red phase for the other lanes;

2 (b): green phase for lanes 1, 5 and 6, and red phase for all the other lanes;

3 (c): green phase for lanes 3 and 7, and red phase for the other lanes;

4 (d): green phase for all the lanes of arm n.1 and a red phase for the other lanes.



(a) (b) (c) (d)

Green-red phase duration: 10s Yellow phase duration: 4s

Reward

The following waiting times are defined:

- waiting time *wt* is the time a vehicle spends with speed 0m/s;
- waiting time of vehicles with low priority: $W_{p_1} = 1 \sum wt_{p_1}$
- waiting time of vehicles with intermediate priority: $W_{p_2} = 190 \sum wt_{p_2}$
- waiting time of vehicles with high priority: $W_{p_2} = 200 \sum wt_{p_3}$

The reward function $\boldsymbol{r_t}$ is defined as follows:

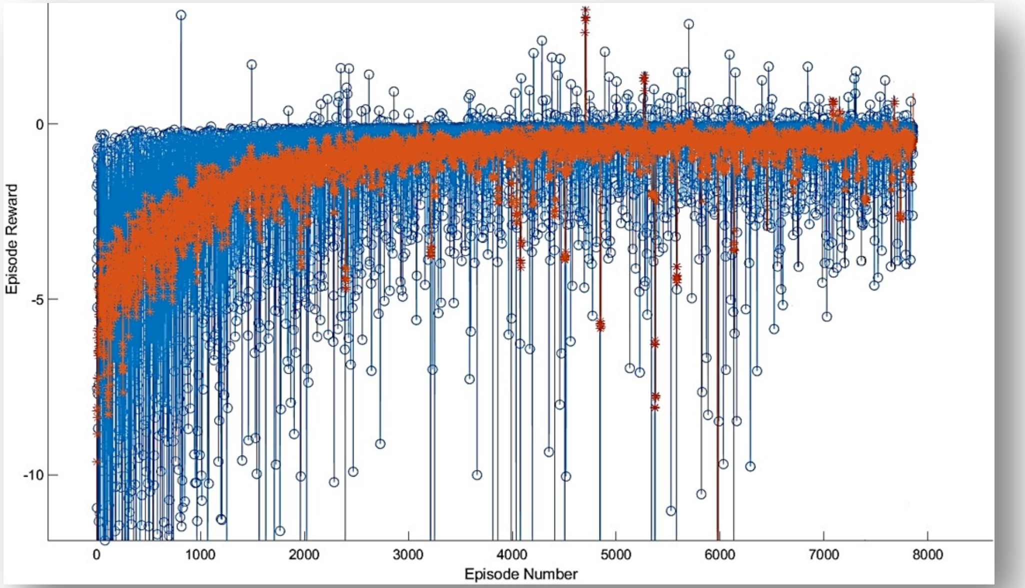
$$r_t = \begin{cases} W_{p_{i_t},t} - W_{p_{i_t},t+1} & if \quad W_{p_{i_t},t} \geq W_{p_{i_t},t+1} \\ 2[W_{p_{i_t},t} - W_{p_{i_t},t+1}] & if \quad W_{p_{i_t},t} < W_{p_{i_t},t+1} \end{cases}.$$

Training

The agent is trained for a total of 8000 episodes and each episode consists of a simulation of a predefined traffic configuration. Vehicles are spawned into each simulation within 1000 seconds from the start.

Number of configuration	Percentages of priorities	Average seconds between each spawn
1	98% P_1, 1,5% P_2, 0,5% P_3	1
2	98% P_1, 1,5% P_2, 0,5% P_3	2
3	98% P_1, 1,5% P_2, 0,5% P_3	3,5
4	98% P_1, 1,5% P_2, 0,5% P_3	5
5	97% P_1, 2% P_2, 1% P_3	1
6	97% P_1, 2% P_2, 1% P_3	2
7	97% P_1, 2% P_2, 1% P_3	3,5
8	97% P_1, 2% P_2, 1% P_3	5

Episode ending condition: All scheduled vehicles leave the simulation or number of time-steps reaches 3000.



Results

The performance of the trained agent when it is facing different traffic situations is tested through the measurement of the average waiting times of the three classes of vehicles.
The performance of a traditional static traffic light is measured in order to compare the two approaches.

Scenario	P_1 average waiting time (seconds)	P_2 average waiting time (seconds)	P_3 average waiting time (seconds)
Traditional traffic light Green-Red: 30s Yellow: 4s 100% P_1	35,41	-	-
Traditional traffic light Green-Red: 10s Yellow: 4s 100% P_1	18,02	-	-
DRL agent 100% P_1	15,58	-	-
DRL agent 98% P_1, 1,5% P_2, 0,5% P_3	18,34	2,22	1,60
DRL agent 97% P_1, 2% P_2, 1% P_3	18,92	2,59	1,74

OUTLINE

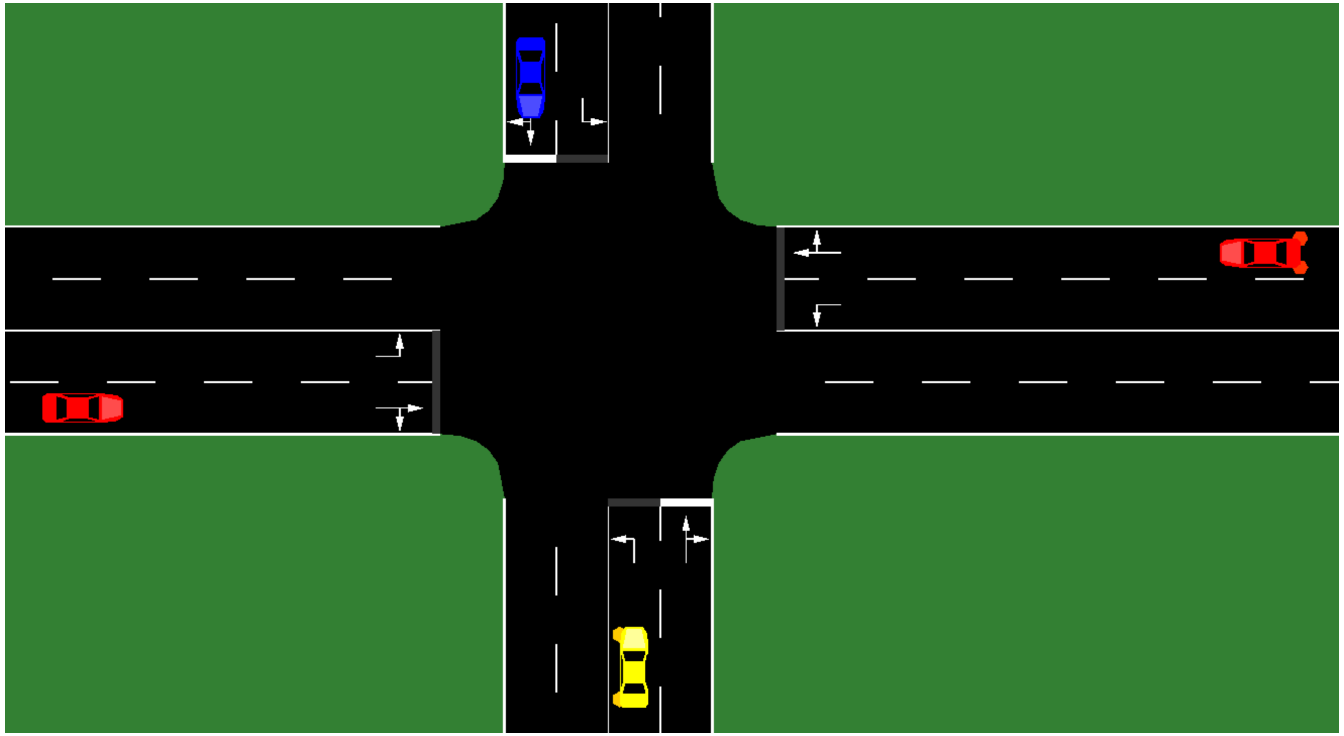
Machine learning (ML) and deep reinforcement learning (DRL) techniques applied in autonomous vehicle control:

- **traffic management** (efficiently handling by intelligent traffic lights road traffic in road intersections where priority issues are important)
- **traffic management** (crossroads without traffic lights involving connected automated vehicles, priority vehicles and regular vehicles)
- **automotive** (autonomous braking systems based on an intelligent agent)

Problem Statement

3 classes of vehicles are approaching a 4-way double lane of crossroads without traffic lights

- red vehicles = autonomos vehicles (CAV)
- blue vehicles = priority vehicles (ambulances) (CPV)
- yellow vehicles = regular vehicles (RV)
- CAVs and CPVs are connected to Infrastructure architecture
- **Exchanged data:** route, current lane, speed, distance to the intersection, distance from the infront vehicle.



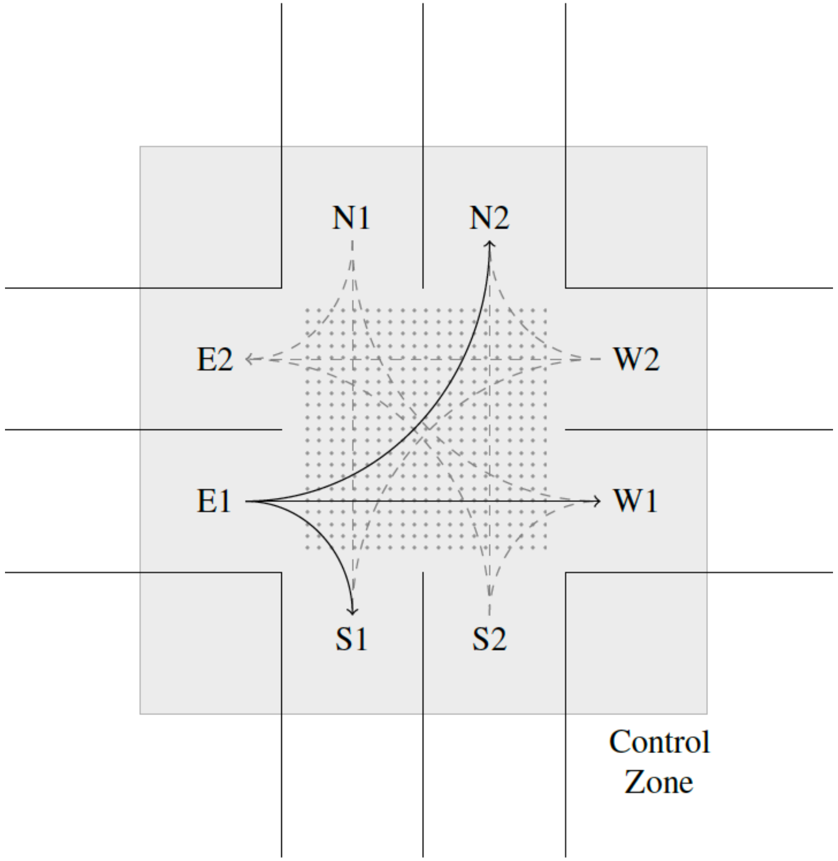
SUMO
SIMULATION OF URBAN MOBILITY

Problem Statement

Objective: design a control strategy for CAVs and CPV that regulates their speed at discrete time intervals to **observe priorities**, **prevent collisions and reduce crossing times**.

The intersection is divided in two zones:

- **control zone** where CAV and CPV communicates with the centralized infrastructure
- **merging zone** where the collisions in the intersection may occur.

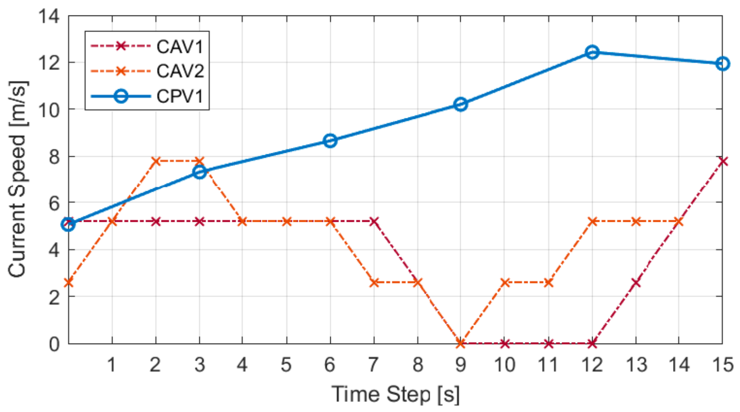


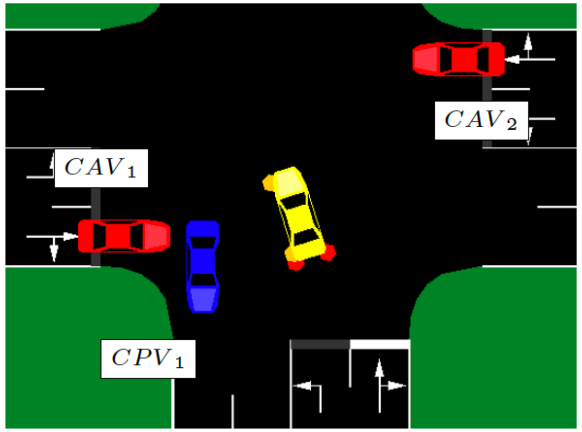
- Agents are the CAVs and the current state of each agent at each time step is designed on the basis of real-time traffic data shared by connected vehicles
- Priority vehicles are connected but driven by humans
- Regular vehicles are unconnected

The process is partially observable.

SUMO
SIMULATION OF URBAN MOBILITY

Reward function and Training



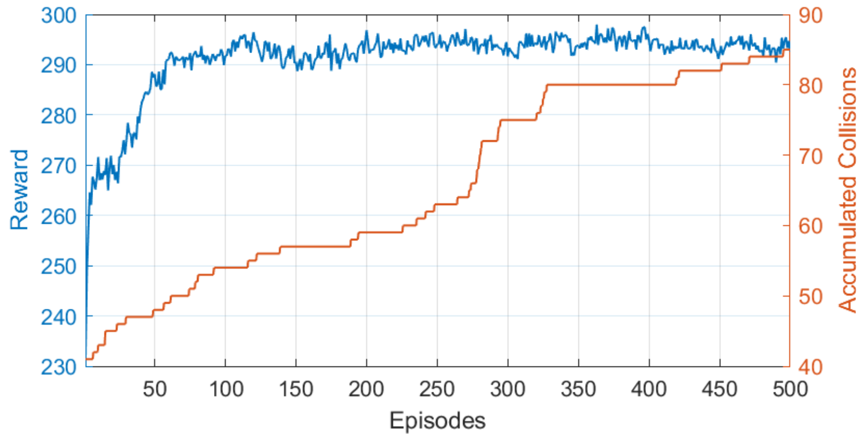


$$p_i(t) = \begin{cases} 1, & \text{if } CAV_i \text{ detects collisions} \\ & \text{with other vehicles at time step } t \\ 0, & \text{otherwise.} \end{cases}$$

Then, the global reward function is defined by:

$$r(t) = w_1 \sum_{i=1}^{n} v_i(t) + w_2 \sum_{j=1}^{m} \hat{v_j}(t) - w_3 \sum_{i=1}^{n} v_i(t) p_i(t),$$

where $w_1, w_2, w_3 \in [0, 1]$ are arbitrary weights.



OUTLINE

Machine learning (ML) and deep reinforcement learning (DRL) techniques applied in autonomous vehicle control:

- **traffic management** (efficiently handling by intelligent traffic lights road traffic in road intersections where priority issues are important)
- **traffic management** (crossroads without traffic lights involving connected automated vehicles, priority vehicles and regular vehicles)
- **automotive** (autonomous braking systems based on an intelligent agent)

Problem statement

An intelligent autonomous braking system enabling to locate pedestrians crossing the roadway and thus stop the vehicle to allow pedestrians to cross.

The system is controlled by a DRL agent aiming at performing **safe and comfortable braking actions**.

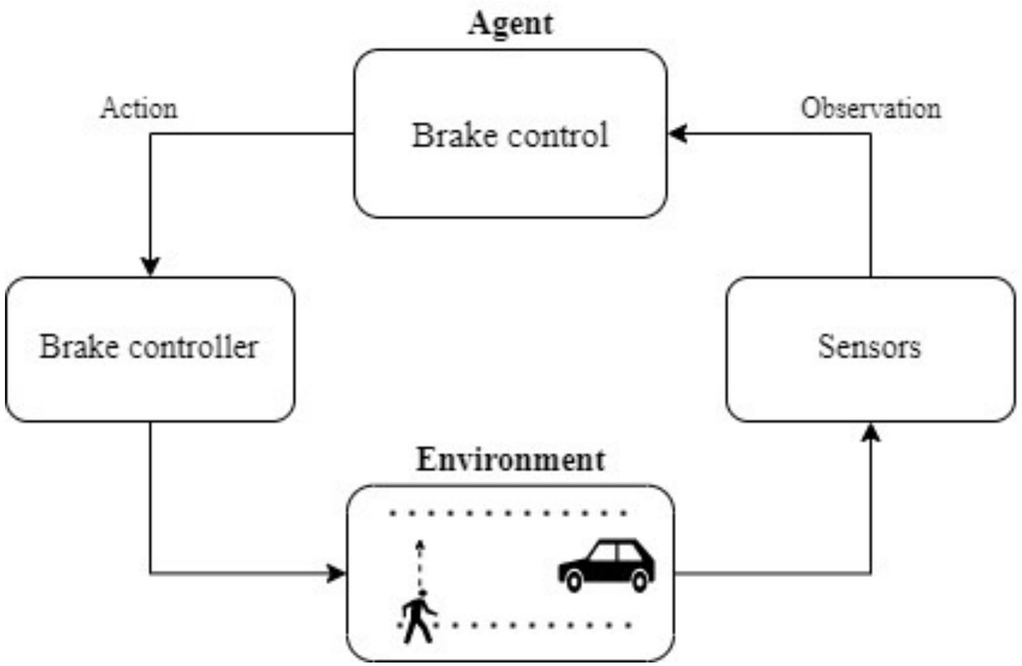
The vehicle braking system is driven by an intelligent agent that

- monitors the presence of crossing pedestrians at each time step
- decelerates considering its speed and pedestrians' distance
- optimizes passengers' comfort by avoiding abrupt actions

Case Study

The environment: a vehicle travelling on a road with pedestrian crossings.

The intelligent agent
- perceives the state of the environment
- applies a braking action to avoid collision
- gets a reward that measures the ”goodness” of its action.

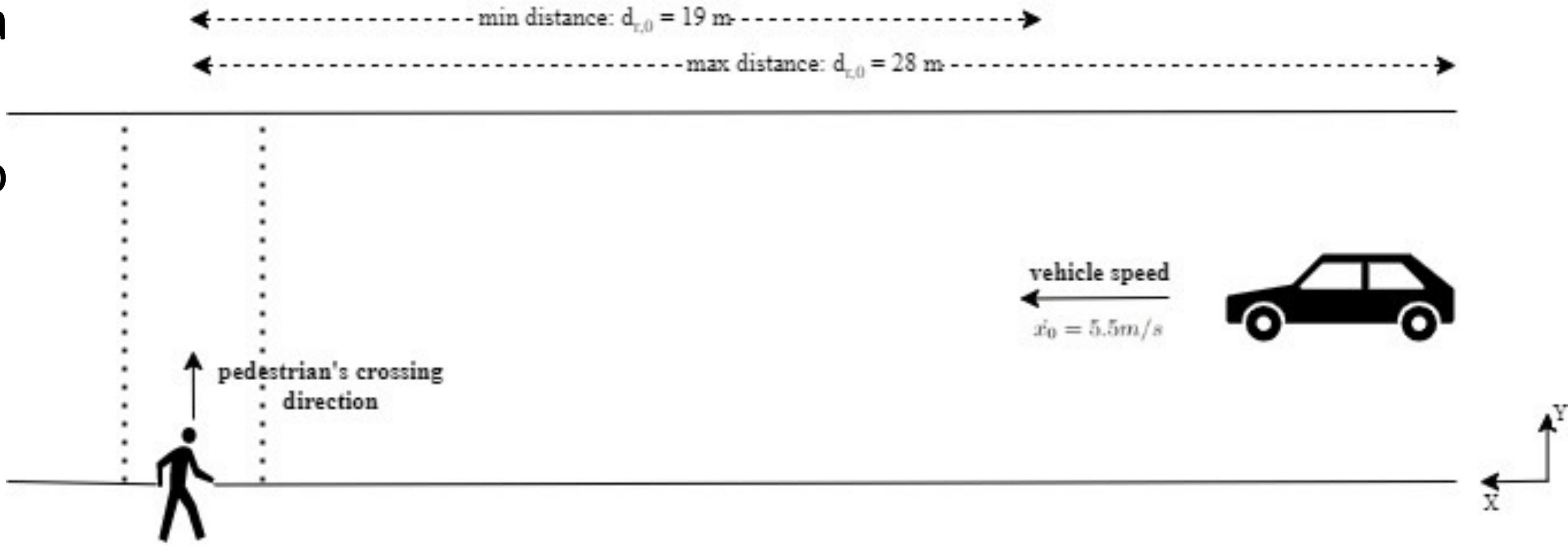


Case Study

- **Training scenario:** the vehicle ca move only in the X direction.
- Pedestrians can cross the road o perpendicularly to the vehicle direction.

In each training phase episode:

- Car initial speed: $\dot{x_0}$= 5.5m/s
- Initial distance between car and pedestrian: random $d_{r,0} \in [19, 28]$m.
- Time step: T_s = 0.1s
- Episode duration: T_f = 20s
- Maximum number of steps per episode: 200.



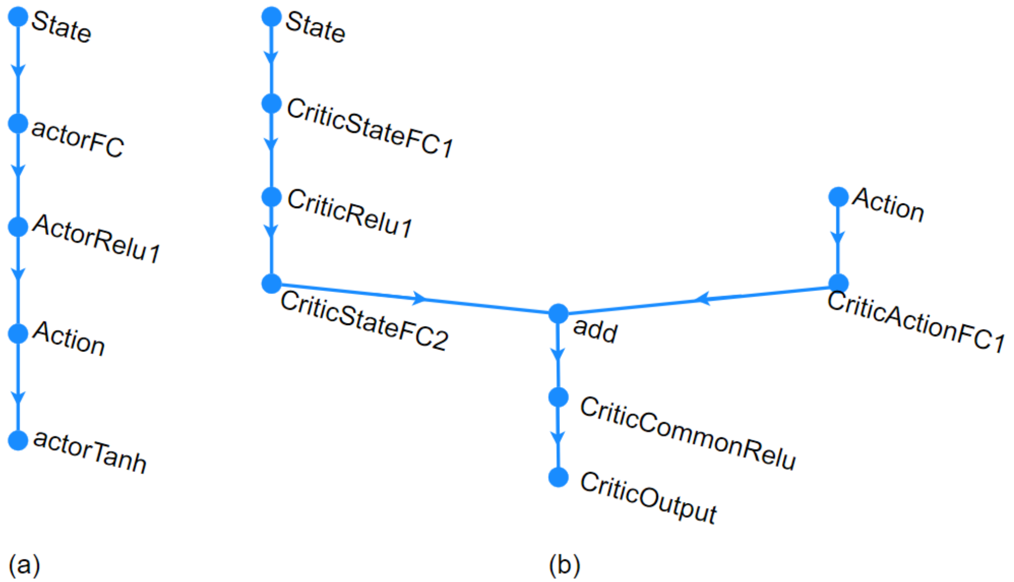
DRL MODEL

A. State Space

- car position x_t
- car speed $\dot{x}_t$
- car braking action in term of deceleration b_t
- position y_t of the pedestrian crossing the road
- distance between car and pedestrian $d_{r,t}$,

At each time step, the agent performs the braking action $b_t \in A = [0, 3]$ m/s^2.

The system is modelled in Matlab and Simulink software environment.



Neural network architecture

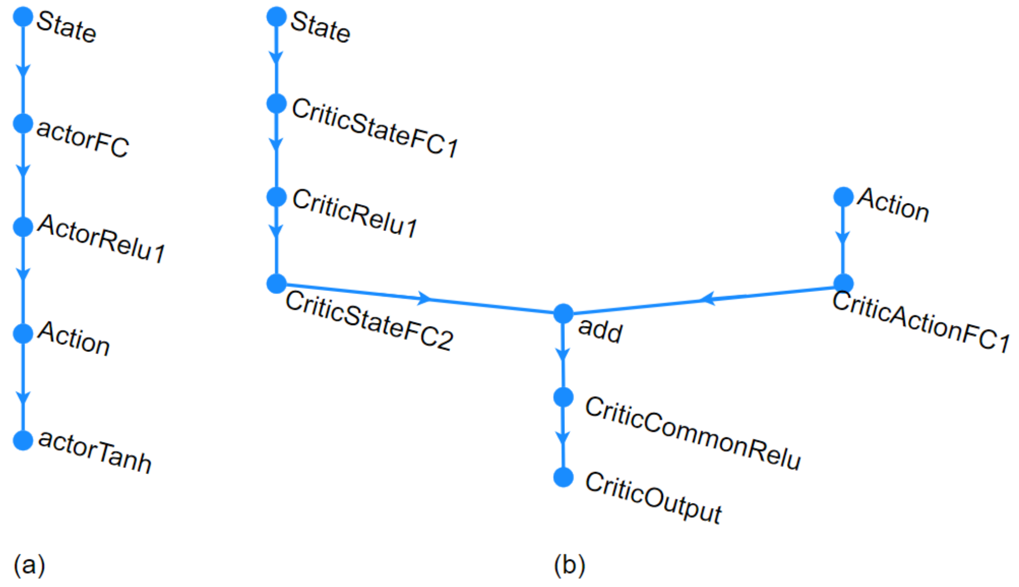
DRL MODEL

A. State Space

- car position x_t
- car speed $\dot{x}_t$
- car braking action in term of deceleration b_t
- position y_t of the pedestrian crossing the road
- distance between car and pedestrian $d_{r,t}$,

At each time step, the agent performs the braking action $b_t \in A = [0, 3]$ m/s².

The system is modelled in Matlab and Simulink software environment.



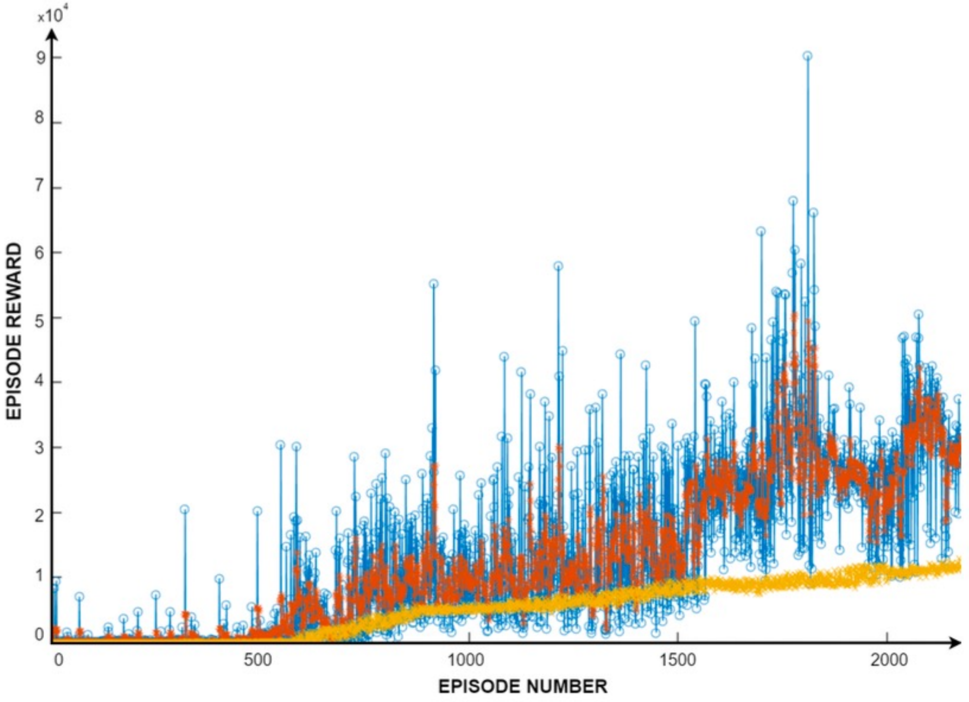
Neural network architecture

Reward strategy: discrete reward

$$r_t = r_1 \cdot r_2$$

$$r_1 = \begin{cases} 50 \text{ if } d_{r,t} > d_s \text{ and } d_{r,t} \leq 2.5 \\ 40 \text{ if } d_{r,t} > 2.5 \text{ and } d_{r,t} \leq 3 \\ 35 \text{ if } d_{r,t} > 3 \text{ and } d_{r,t} \leq 3.5 \\ 30 \text{ if } d_{r,t} > 3.5 \text{ and } d_{r,t} \leq 4 \\ 20 \text{ if } d_{r,t} > 4 \text{ and } d_{r,t} \leq 4.5 \\ 10 \text{ if } d_{r,t} > 4.5 \text{ and } d_{r,t} \leq 5.5 \\ 1 \text{ if } d_{r,t} > 5.5 \text{ and } d_{r,t} \leq 15 \\ 0 \text{ if } d_{r,t} < d_s \text{ and } d_{r,t} > 15 \end{cases}$$

$$r_2 = \begin{cases} 5 \text{ if } b_t > 0 \text{ and } b_t \leq 1.5 \\ 3 \text{ if } b_t > 1.5 \text{ and } b_t \leq 2 \\ 2 \text{ if } b_t > 2 \text{ and } b_t \leq 2.5 \\ 1 \text{ if } b_t > 2.5 \text{ and } b_t \leq 3 \end{cases}$$



The agent gets fixed amount of reward depending on observations

Agent tests

The reward strategy shows effectiveness and efficiency allowing the vehicle to stop by comfortable and continuous braking and precisely at a suitable distance from the pedestrian.

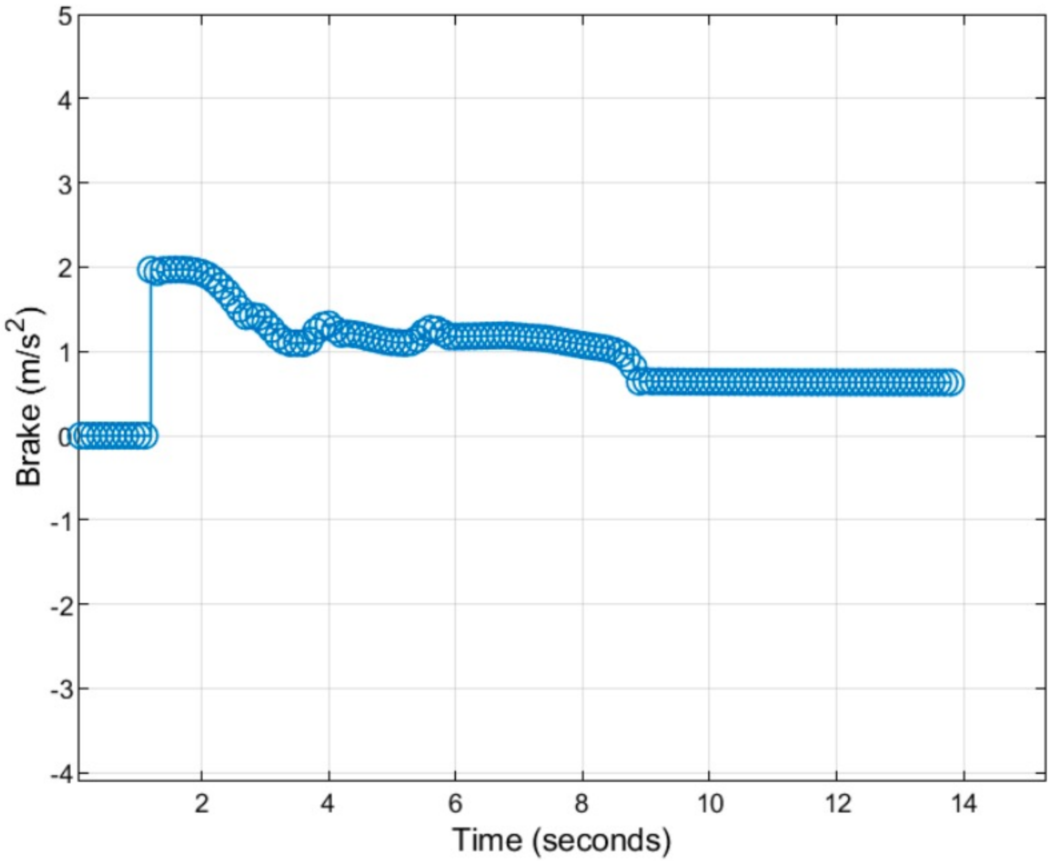


Fig. 5: Agent's test with discrete reward strategy: braking diagram.

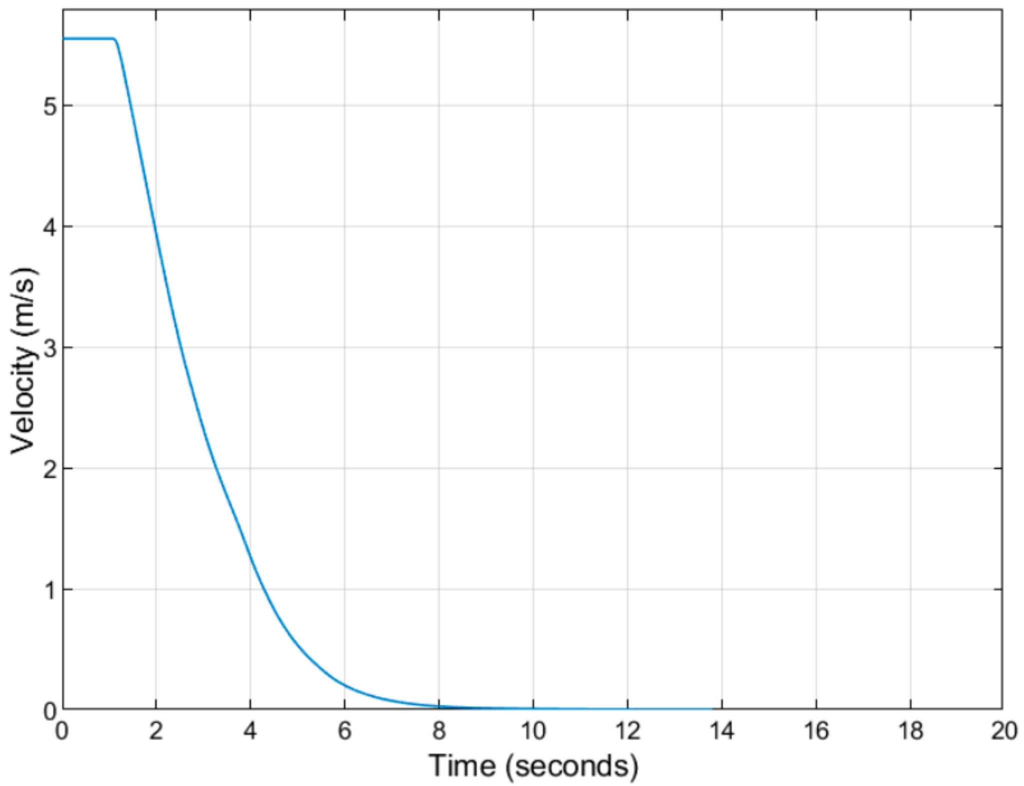


Fig. 6: Agent's test with discrete reward strategy: speed diagram.

Conclusions

DIPARTIMENTO DI INGEGNERIA ELETTRICA E DELL'INFORMAZIONE

Machine learning and Deep Reinforcement Learning are efficiently applied in intelligent transportation systems and promising results are expected.

- **traffic management** (efficiently handling by intelligent traffic lights road traffic in road intersections where priority issues are important)
- **traffic management** (crossroads without traffic lights involving connected automated vehicles, priority vehicles and regular vehicles)
- **automotive** (autonomous braking systems based on an intelligent agent)

G. Benedetti, M. P. Fanti, A. M. Mangini and F. Parisi, "Application of Deep Reinforcement Learning for Traffic Control of Road Intersection with Emergency Vehicles," *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2021, pp. 182-187.

M. P. Fanti, A. M. Mangini, D. Martino, I. Olivieri, F. Parisi, F. Popolizio "Safety and Comfort in Autonomous Braking System with Deep Reinforcement Learning", *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2022, .

G. Volpe, M. P. Fanti, A. M. Mangini, "A Cooperative DRL Approach for Autonomous Traffic Prioritization in Mixed Vehicles Scenarios" *2023 IEEE International Conference on Atomation and Systems Engineering)*, accepted at IEEE CASE 2023.

ICRA

LONDON · 2023

EMBRACING THE FUTURE: MAKING ROBOTS FOR HUMANS

29 May – 2 June 2023 | ExCeL London

ICRA Automation Cluster Forum

Thanks for your attention!

Contact: **mariapia.fanti@poliba.it**

Politecnico di Bari

DIPARTIMENTO DI INGEGNERIA ELETTRICA E DELL'INFORMAZIONE

LCA