HyperDrive: Autonomous Self-Driving Car in an Urban Setting using Deep Reinforcement Learning

Sana Haider
p180011@nu.edu.pk

Syed Asad Zaman
p180034@nu.edu.pk

Supervisor
Mr. Waqas Ali

November 10, 2021

Literarture Review

Literarture Review

Year	Title	Achievements	Limitations
2015	Human-level control through deep reinforcement learning [1][1]	Developed a deep Q-network,that can learn successful policies directly from high-dimensional sensory inputs using end-to-end reinforcement learning	DQN is limited to 2D Games only with limited action & Search Space
2016	Mastering the game of go without human knowledge [2][2]	Developed a GO Gaming agent using Deep Q-Network (DQN) and Monte-Carlo tree search	The agent was taking too much time to make decisions

[1] Mnih, Volodymyr and Kavukcuoglu, Koray and Silver, David and Rusu, Andrei A and Veness, Joel and Bellemare, and others, **Human-level control through deep reinforcement learning**, Nature Publishing Group, 2015

[2] Silver, David and Schrittwieser, Julian and Simonyan, Karen and Antonoglou, Ioannis and Huang, Aja and Guez, Arthur and Hubert, , **Mastering the game of go without human knowledge**, Nature Publishing Group, 2017

Literature Review

Year	Title	Achievements	Limitations
2017	CARLA: An open urban driving simulator[3] [3]	An Open Source Urban Driving Simulator using Unreal Engine 4	——
2018	Self-driving cars using CNN and Q-learning [4] [4]	An autonomous Self-Driving Car from Scratch which was able to self-drive in constraint environment	Environment was static & limited only to traffic sign and self made path

[3] Dosovitskiy, Alexey and Ros, German and Codevilla, Felipe and Lopez, Antonio and Koltun, Vladlen ,**CARLA: An open urban driving simulator**, Conference on robot learning, 2017

[4] Chishti, Syed OwaisAli and Riaz, Sana and BilalZaib, Muhammad and Nauman, Mohammad, **Self-driving cars using CNN and Q-learning**, 2018 IEEE 21st International Multi-Topic Conference (INMIC), 2018

Literature Review

Year	Title	Achievements	Limitations
2020	Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment [5][5]	This paper proposed a two- layered path planning model, method includes a high-level model that produces a rough path and a low-level model that provides precise navigation	High number of sampling doesn't give the best solution
2021	An Adaptive Ant Colony Algorithm for Autonomous Vehicles Global Path Planning [6][6]	This paper proposed path planning in a obstacle oriented environment using the optimized version of Ant Colony Algorithm	Dynamic Path Planning is not discussed in this paper

[5] Chen, Jiajia and Zhang, Rui and Han, Wei and Jiang, Wuhua and Hu, Jinfang and Lu, Xiaoshan and Liu, Xingtao and Zhao, Pan,**Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment**, Journal of Advanced Transportation, 2020

[6] Li, Yanqiang and Ming, Yu and Zhang, Zihui and Yan, Weiqi and Wang, Kang, **An Adaptive Ant Colony Algorithm for Autonomous Vehicles Global Path Planning**, 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2021

Literature Review

Year	Title	Achievements	Limitations
2021	Path planning for autonomous ground vehicles based on quintic trigonometric Bézier curve [7][7]	This paper proposed the method of Predefined Path planning using Bézier curve	Role of Dynamic Actors are not discussed in this paper
2018	Path planning algorithm using D* heuristic method based on PSO in dynamic environment [8][8]	This paper proposed method Shortest safe path planning using D* algorithm based on Particle Swarm Optimization (PSO)	The method is only applicable to Limited static Environment

[7] Bulut, Vahide, **Path planning for autonomous ground vehicles based on quintic trigonometric Bézier curve**, Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2021

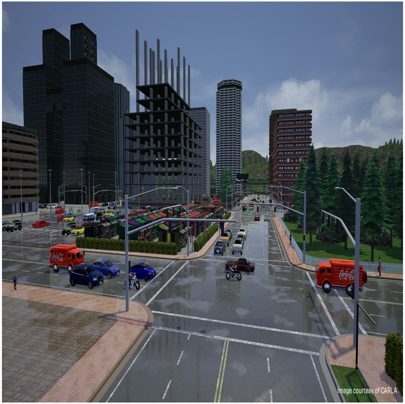
[8] Raheem, Firas A and Hameed, Umniah I and others,**Path planning algorithm using D* heuristic method based on PSO in dynamic environment**, American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 2018

Problem Statement

Problem Statement

Autonomous Self-Driving Car's Routing in a Dynamic Urban Environment of CARLA.

Urban Environment of CARLA



UML Diagrams

Use Case Diagram

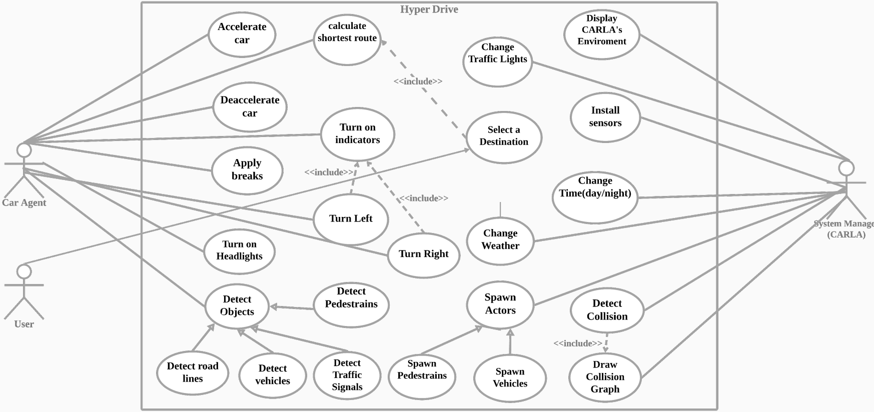


Figure 1: Use Case Diagram

Activity Diagram

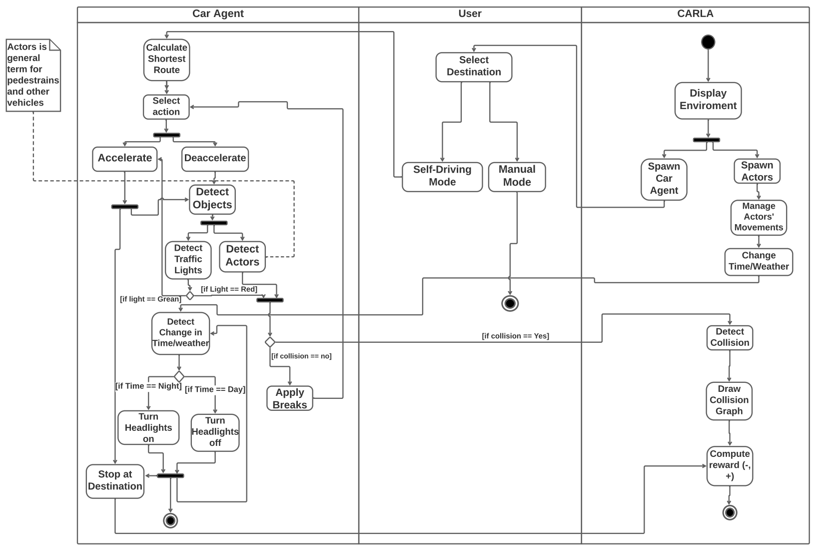


Figure 2: Activity Diagram

Proposed Solution

Proposed Solution

D* Search[1]

[1] Raheem, Firas A and Hameed, Umniah I and others,**Path planning algorithm using D* heuristic method based on PSO in dynamic environment**, American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 2018

[2] Mnih, Volodymyr and Kavukcuoglu, Koray and Silver, David and Rusu, Andrei A and Veness, Joel and Bellemare, and others, **Human-level control through deep reinforcement learning**, Nature Publishing Group, 2015

Proposed Solution

D* Search[1]

Deep Q-Network(DQN)[2]

[1] Raheem, Firas A and Hameed, Umniah I and others,**Path planning algorithm using D* heuristic method based on PSO in dynamic environment**, American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 2018

[2] Mnih, Volodymyr and Kavukcuoglu, Koray and Silver, David and Rusu, Andrei A and Veness, Joel and Bellemare, and others, **Human-level control through deep reinforcement learning**, Nature Publishing Group, 2015

Proposed Solution

D* Search[1]

Deep Q-Network(DQN)[2]

Dual-Deep Q-Network

[1] Raheem, Firas A and Hameed, Umniah I and others,**Path planning algorithm using D* heuristic method based on PSO in dynamic environment**, American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 2018

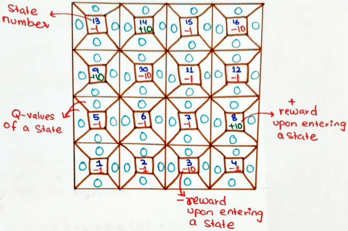
[2] Mnih, Volodymyr and Kavukcuoglu, Koray and Silver, David and Rusu, Andrei A and Veness, Joel and Bellemare, and others, **Human-level control through deep reinforcement learning**, Nature Publishing Group, 2015

Mathematical Equation

Bellman Equation

$$\underbrace{Q(s,a)}_{\substack{\text{New}\\ \text{Q-Value}}} = \underbrace{Q(s,a)}_{\substack{\text{Current}\\ \text{Q-Value}}} + \underbrace{\alpha}_{\substack{\text{Learning}\\ \text{rate}}} \Big[\underbrace{R(s,a)}_{\text{Reward}} + \underbrace{\gamma}_{\substack{\text{Discounting}\\ \text{factor}}} \overbrace{\max Q'(s',a')}^{\substack{\text{Maximum predicted reward, given}\\ \text{new state and all possible actions}}} - Q(s,a) \Big]$$

Working



▶ **Starting Situation:**

initially ;

All Q-values = 0

α = 0.8

γ = 0.6

Randomly

Initial State = 1

$$Q(S_t, a_t) \leftarrow Q(S_t, a_t) + \alpha\left[R(S_t, a_t) + r \cdot \max Q(S_{t+1}, a) - Q(S_t, a_t)\right]$$

→ Agent is in State 1, taking 'Right' Action, it will land in State 2 :

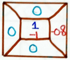
Putting values in equation

$$Q(S_1, \text{Left}) \leftarrow Q(S_1, \text{left}) + \alpha\left[R(S_1, \text{left}) + r \cdot \max Q(S_2, a) - Q(S_1, \text{left})\right]$$

$$= 0 + 0.8(-1 + 0.6(0) - 0)$$

$$Q(S_1, \text{Right}) = -0.8$$

State 1 updated!



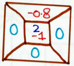
Now, Agent is in State 2 : taking up action will land the agent in State 6:

Putting values in equation:

$$= 0 + 0.8(-1 + 0.6(0) - 0)$$

$$Q(S_2, \text{up}) = -0.8$$

State 2 updated!



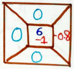
Agent is in State 6, taking Right Action will result in State 7 :

Putting values in equation.

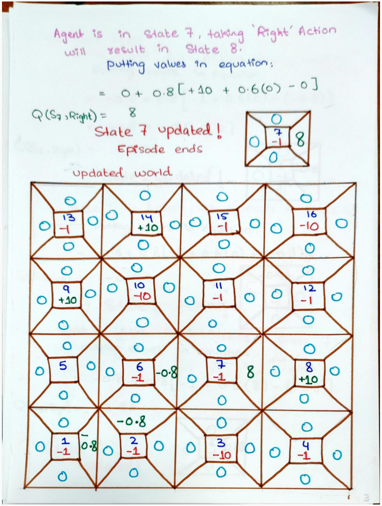
$$= 0 + 0.8(-1 + 0.6(0) - 0)$$

$$Q(S_6, \text{Right}) = -0.8$$

State 6 updated!



Working (cont)


Agent is in State 7, taking 'Right' Action will result in State 8.

Putting values in equation:

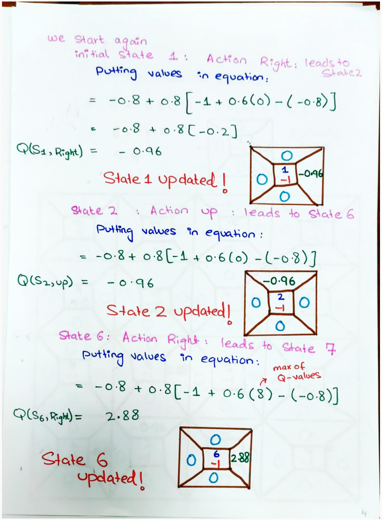
$$= 0 + 0.8[+10 + 0.6(0) - 0]$$

$Q(S_7, Right) = 8$

State 7 updated!

Episode ends

updated world


we start again

initial State 1 : Action Right: leads to State2

Putting values in equation:

$$= -0.8 + 0.8[-1 + 0.6(0) - (-0.8)]$$

$$= -0.8 + 0.8[-0.2]$$

$Q(S_1, Right) = -0.96$

State 1 updated!

State 2 : Action up : leads to State 6

Putting values in equation:

$$= -0.8 + 0.8[-1 + 0.6(0) - (-0.8)]$$

$Q(S_2, up) = -0.96$

State 2 updated!

State 6: Action Right: leads to State 7

putting values in equation:

max of Q-values

$$= -0.8 + 0.8[-1 + 0.6(8) - (-0.8)]$$

$Q(S_6, Right) = 2.88$

State 6 updated!

Working (cont)

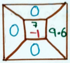
State 7, Action Right, resultant State = 8.

Putting values in equation

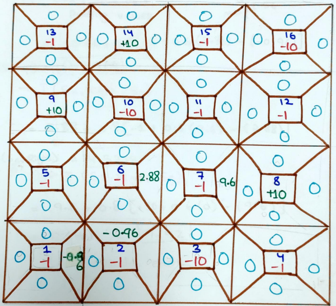
$= 8 + 0.8[10 + 0.6(0) - (8)]$

$= 8 + 1.6$

$Q(S_7, Right) = 9.6$



Episode ends



→ Again, Start from State 1:

initial State : 1

Action : Right → leads to State 2

Putting values

$= -0.96 + 0.8[-1 + 0.6(0) - (-0.96)]$

$Q(S_1, Right) = \boxed{-0.992}$

→ $T(S_2, up, S_6)$

Putting values

Q-value max of S_6

$= -0.96 + 0.8[-1 + 0.6(2.88) - (-0.96)]$

$Q(S_2, up) = \boxed{0.3904}$

→ $T(S_6, Right, S_7)$

$= 2.88 + 0.8[-1 + 0.6(9.6) - (2.88)]$

$Q(S_6, Right) = \boxed{4.5344}$

→ $T(S_7, Right, S_8)$

$= 9.6 + 0.8[+10 + 0.6(0) - (9.6)]$

$Q(S_6, Right) = \boxed{9.92}$

Episode ends.

Working (cont)

Start Again :

initial state 1 :

Action : Right

$T(S_1, \text{Right}, S_2)$:

Putting values in equation :

$$= -0.992 + 0.8[-1 + 0.6(0.3904) - (-0.992)]$$

$$Q(S_1, \text{Right}) = -0.811008$$

$T(S_2, \text{Right}, S_3)$:

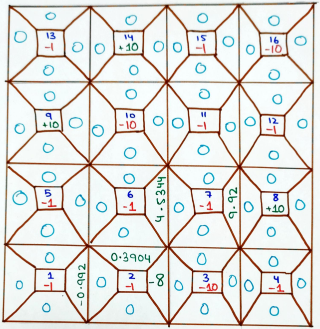
Putting values in equation :

$$= 0 + 0.8[-10 + 0.6(0) - 0]$$

↳ large - Reward

$$Q(S_2, \text{Right}) = -8$$

Episode ends ...



Demo

Demo

CARLA

References

References i

[1] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[2] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton *et al.*, "Mastering the game of go without human knowledge," *nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[3] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on robot learning*. PMLR, 2017, pp. 1–16.

[4] S. O. Chishti, S. Riaz, M. BilalZaib, and M. Nauman, "Self-driving cars using cnn and q-learning," in *2018 IEEE 21st International Multi-Topic Conference (INMIC)*. IEEE, 2018, pp. 1–7.

[5] J. Chen, R. Zhang, W. Han, W. Jiang, J. Hu, X. Lu, X. Liu, and P. Zhao, "Path planning for autonomous vehicle based on a two-layered planning model in complex environment," *Journal of Advanced Transportation*, vol. 2020, 2020.

[6] Y. Li, Y. Ming, Z. Zhang, W. Yan, and K. Wang, "An adaptive ant colony algorithm for autonomous vehicles global path planning," in *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE, 2021, pp. 1117–1122.

[7] V. Bulut, "Path planning for autonomous ground vehicles based on quintic trigonometric bézier curve," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 43, no. 2, pp. 1–14, 2021.

[8] F. A. Raheem, U. I. Hameed *et al.*, "Path planning algorithm using d* heuristic method based on pso in dynamic environment," *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, vol. 49, no. 1, pp. 257–271, 2018.

Thank you!

Any Questions?