BIO Web of Conferences **97**, 00011 (2024) https://doi.org/10.1051/bioconf/20249700011
ISCKU 2024

Dynamic Path Planning using a modification Q-Learning Algorithm for a Mobile Robot

Noor H. Fallooh[1*], *Ahmed* T. Sadiq[2], *Eyad* I. Abbas[3] and *Ivan* A. hashim[4]

[1,3,4]Electrical Engineering Department, University of Technology-Iraq
[2]Computer Science Department, University of Technology-Iraq

Abstract. Robot navigation involves a challenging task: path planning for a mobile robot operating in a changing environment. This work presents an enhanced Q-learning based path planning technique. For mobile robots operating in dynamic environments, an algorithm and a few heuristic searching techniques are suggested. Enhanced Q-learning employs a novel exploration approach that blends Boltzmann and ε-greedy exploration. Heuristic searching techniques are also offered in order to constrict the orientation angle variation range and narrow the search space. In the meantime, the robotics literature of the energy field notes that the decrease in orientation angle and path length is significant. A dynamic reward is suggested to help the mobile robot approach the target location in order to expedite the convergence of the Q-learning and shorten the computation time. There are two sections to the experiments: quick and reassured route planning. With quickly path planning, the mobile robot can reach the objective with the best path length, and with secure path planning, it can avoid obstacles. The superior performance of the suggested strategy is quick and reassured 8-connection Q-learning (Q8CQL) was validated by simulations, comparing it to classical Q-learning and other planning methods in terms of time taken and ideal path.

1 Introduction

Among the key responsibilities in mobile robot navigation these days is path planning. Locating an ideal or nearly ideal collision-free path in relation to certain characteristics in a setting with certain obstacles from a starting place to a goal location [1]. To address the issue of mobile robot path planning, several strategies have been put forth.

Global and local path planning are the two categories into which path planning falls, based on information from the surroundings. The path needs to be found before execution if the environment is a known static terrain. This is known as global path planning. Cell decomposition [2], visibility graphs [3], Artificial potential fields [5] and Voronoi diagrams [4] are examples of global path planning techniques. When the environment is complicated, global path planning techniques are computationally costly. Should the surroundings be partially or completely unknown terrain; local path planning is what it is. Numerous local path planning techniques exist, including those based on fuzzy logic [6], simulated annealing method [7], ant colony algorithms (ASO) [8], evolutionary algorithms [9], particle swarm optimization (PSO) [10], and reinforcement learning (RL).

Among the various approaches currently in use to address the path planning problem, reinforcement learning (RL) stands out for having an alternate learning mechanism that is based on the idea of rewards and penalties. One of the most well-known RL algorithms is Q-learning, which has been successfully used in a variety of applications, including multi-agent intelligent decision-making, vehicle dispatch, and mobile robot control. It is also applied to the problem of mobile robot path planning. Goswami and associates [11].

suggested an extended Q-learning (EQL) model, where the best action at a state is the only one stored in the Q-table. Planning is not possible when the optimal course of action presents a challenge. An enhanced Q-learning was proposed by Amit et al. [13] based on the research in [12]. Unlike the traditional Q-learning (CQL), which updates the Q-values

*Corresponding author: Eee.21.14@grad.uotechnology.edu.iq

periodically, enhanced Q-learning does not update the Q-value when the state is locked. In [14], a heuristic Q-learning algorithm for mobile robot navigation is presented; nevertheless, it is limited to application in basic static environments. A meme method that employed Q-learning and differential evolution to address multi-robot path planning was presented by Pratyusha et al. [15]. The mobile robot path planning problem has been solved for a long time using the Q-learning method, but it is still a challenging challenge. a dynamic setting using the Q-learning model. The path planning problem for mobile robots in a dynamic environment is addressed in this study using a new approach based on an enhanced Q-learning algorithm and a few heuristic searching techniques. Boltzmann and ε-greedy exploration are combined in enhanced Q-learning to create a novel exploration technique for choosing action at states. In contrast to the conventional ε-greedy approach, the new strategy can appropriately balance the tradeoffs between exploration and exploitation. Furthermore, in order to meet the heuristic searching techniques, random action must be carried out throughout the exploration phase. The search space for enhanced Q-learning is smaller than CQL after the heuristic searching algorithms are introduced. Consequently, the suggested approach is successful in preventing regional as compared to the CQL algorithm, and to minimize the time required. Additionally, as compared to CQL, the robot's orientation angle is lower. There are experiments that support the effectiveness of the suggested approach. Every one of the aforementioned algorithms has advantages and disadvantages. The majority of the studies focus on exact modeling and location navigation depending on the surroundings [16]. However, little to nothing is known about the actual ecosystems. It calls for Mobile robot avoidance of unidentified challenges that are both static and dynamic. Furthermore, the majority of publications use Mobile robot as a mass point for path planning, which causes the path to be excessively near to barriers and causes Mobile robot to collide with them [17].

The suggested method allows the mobile robot to plan the course in two modes: quick path planning and safe path planning, which addresses the shortcomings of the previous path planning technique. Simulations and comparative trials are used to illustrate the quick and reassured 8-connection Q-learning (Q8CQL) algorithm's superiority, efficacy, and speed. The following primary contributions are made when the (Q8CQL) technique presented in this research is used to solve the path planning for mobile robots that compared to motion planning suggestions based on the most recent algorithms, the Q8CQL is a novel proposal that can identify short and safe conditions for Mobile robot route planning. There are two suggested path planning schemes: the first plans the secure path length (far from barriers) from the starting point to the destination point, while the other plans the shortest path length (close to obstacles). The suggested method avoids a dead-end path obstructed by obstacles and accelerates the convergence of the algorithm by converting the constant reward into a blend of static and dynamic reward.

The suggested approach in this paper is shown to be able to solve the mobile robot path planning problems for both short and safe path planning modes by comparison with other algorithms.

2 Related work

Siding Li, Xin Xu, Lei Zuo in [2015] [18]. This research proposes a novel approach to path planning for mobile robots operating in dynamic environments. It is based on an enhanced Q-learning (IQL) algorithm and a few heuristic searching techniques. IQL employs a novel exploration approach that blends Boltzmann and ε-greedy exploration. Moreover, heuristic searching techniques are offered to constrict the orientation angle variation range and narrow the search space. The suggested method's superior performance was validated using simulations, comparing it to classical Q-learning (CQL) and other planning approaches in terms of time taken and ideal path. The robotics literature on energy consumption, however, places importance on the reduction of orientation angle and path length.

Ee Soong Low, Pauline Ong, Kah Chun Cheah in 2019 [19]: The concept of partially guided Q-learning is introduced to get around the limitation. In this method, the flower pollination algorithm (FPA) is used to improve the Q-learning initialization. When Q-values are initiallyized correctly using the FPA, Q-learning can accelerate its convergence, as demonstrated by the experimental evaluation of the enhanced Q-learning that has been proposed in the challenging environment utilizing an alternative barrier configuration.

MONIRA ESSA ALOUD ANDNORAALKHAMEES in 2021 [20]: When Reinforcement Learning (RL) is applied to price time-series data, it can attain optimal dynamic algorithmic trading. It is true that proposing a dynamic algorithmic trading utilizing reinforcement learning requires a thorough description of the environment states. As a result, we suggest employing the Directional Change (DC) event technique along with a dynamic DC threshold to describe the environmental states. The suggested algorithmic trading strategy is referred to as the DCRL trading strategy. Furthermore, an optimal trading rule and the Q-learning algorithm were used to train the suggested DCRL trading strategy. Using actual stock market data, we assessed the DCRL trading strategy. The findings show that, in a turbulent stock market, the DCRL state representation policies increased trading returns and enhanced Sharpe Ratios.

Talal Bonny and Mariam Kashkash in 2021[21]:The objective of hybridizing Q-Learning and BA is to determine the optimal path with the fewest number of BA iterations. The shortest path is found using the Q-Learning sterilizing feature, and the problem can be solved without any limitations by using the BA. The experiment is carried out on several different maps in order to validate the proposed method in both the static and the dynamic instances. The results of the experiment show how stable and effective the recommended approach is in figuring out what to do.

He Du, Bing Hao and et in 2022 [22] An original method termed short and safe Q-learning is described in the paper that simplifies the short and safe path planning problem for mobile robots. To mitigate the delayed convergence of Q-learning, the artificial potential field is employed to inhibit haphazard exploration and provide mobile robots with a priori awareness of their environment. To further accelerate the convergence of the Q-learning and reduce computation time, a dynamic incentive is proposed to assist the mobile robot in approaching the goal position. Safe path planning and short path planning are the two sections that comprise the experiments. A mobile robot that uses safe route planning can avoid obstacles while using short path planning to go to the target with the shortest possible path length.

3 Materials and methods

3.1 Problem formulation of mobile robot

The path planning method is depicted in Fig. 1 to help better demonstrate the planning challenge using Q8CQL in this paper. The definition of safe distance and the schematic diagram of multiple regression are displayed in Figure 1a, where r stands for the circular mobile robot's radius and xa and ya for the earth-fixed frame and xb and yb for the body-fixed frame [23]. It is believed that the MR center of mass is where the body-fixed frame originated. This study defines the MR safety radius as rsd to help plan the path away from obstacles and prevent the planned path from colliding with them. This paper defines to improve obstacle avoidance when there are no obstacles around mobile robot can move to the adjacent eight directions (down, down left, down right, left, up, up left, up down, right)

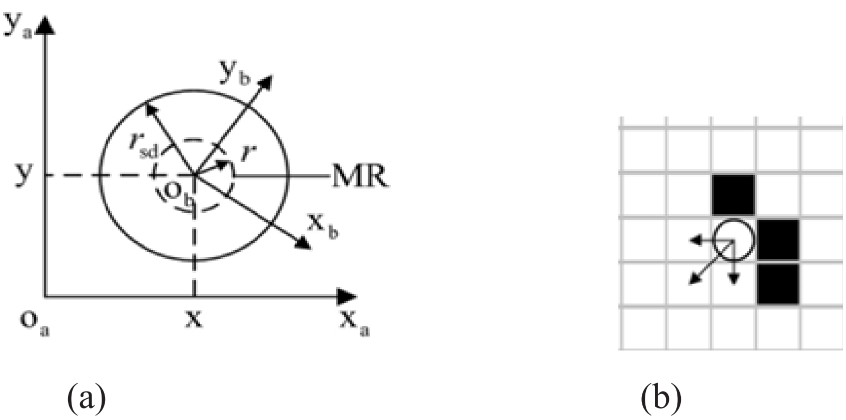


Fig.1. Show problem of mobile robot

The obstacle barrier is too big for mobile robot to navigate diagonally. In Fig. 1b, the mobile robot's movement directions are displayed when the obstacles are situated around it. From the start point S to the destination point T in a Cartesian coordinate, the mobile robot shows an unstoppable viable path free of obstacles [5].

3.2 Q-learning algorithm

Q-learning is a class of reinforcement learning (RL) algorithms that Watkins developed in 1989 [4]. Q-learning uses the concepts of reward and penalty to explore the unstructured environment. The terms state, action, agent, reward, and penalty are used in the Q-learning method, the following equation shows the agent as the mobile robot; an action is the motion an agent performs to change from one state to another, and a state is the location the agent defines in the environment [24]. When an agent performs the right action at a given state, a reward is a positive value that raises the Q-value, while a penalty is a negative value that lowers the Q-value.

$$Q(s_t, a_t) \leftarrow (1-\alpha) * Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \, max_a \, Q(s_{t+1}, a_{t+1})] \quad (1)$$

The equation (1) can also be written as equation (2) and (3)

$$Q(s_t, a_t) \leftarrow (1-\alpha) * Q(s_t, a_t) + \alpha \, [r_{t+1} + \gamma \, V(s_{t+1})] \quad (2)$$

$$V(s_t) \leftarrow V(s_t) + \alpha[r_{t+1} + \gamma \, V(s_{t+1}) - V(s_t) \quad (3)$$

Where: s_t :current state, a_t :Action performed in s_t state, r_{t+1}: Received reinforcement signal after s_t executed, s_{t+1} :Next state, α: learning coefficient (0<α<1) and γ: discount factor (0<γ<1).

For the following reasons, mobile robot path planning uses Q-learning:

1. Reinforcement learning interacts well with the environment [25]. By investigating and learning from their surroundings, mobile robot expands their present knowledge and refines their operating tactics to better context.

2. As time permits, the Q-learning algorithm makes several attempts at achieving the optimal mal policy for every potential pair of state actions [26]. It is an iterative, highly experimental process.

3. Q-learning makes advantage of off-policy [27], meaning that the distance between mobile robot and the obstacle can be controlled by choosing actions based on the target strategy [28].

A quick and reassured 8-connection approach called Q8CQL is proposed to solve the problem of secure and speedy path planning. The mobile robot path planning problem is a better way to represent the study's findings.

In reality, both the obstacle environment and the mobile robot are three-dimensional entities. Their motion space is seen as two-dimensional coordinates in this research to simplify the problem.

1. Short and secure path lengths comprise the two components of this paper's mobile robot route planning research. Environmental obstacles are considered in planning the path, taking into account their size, shape, and location.

2. In order to immediately reach the objective, the mobile robot store seeks the shortest and safest path length.

3.3 The Q-table of suggested Q8CQL learning algorithm

During the exploration phase, the Q-table is often initialized to zero or regularly distributed random values using classical Q-learning, which causes the mobile robot to choose interactions at random. resulting in the algorithm's high computation time and delayed convergence. To optimization this problem suggested quick and secure Q-learning algorithm.

By adding the four directions (up, left, down, and right) for movement possibilities at the least angle at which the mobile robot may move without interacting with the surrounding obstacles.

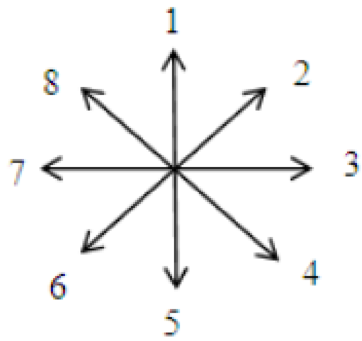


Fig.2. The 8-connection of Q8CQL

the proposed method gives the mobile robot eight connections instead of four. Additionally, the shortest distance between two places is represented by movement in the **inclined** direction with the smallest angle in a cartesian coordinate system is used to represent a grid of environment, with the x- and y-axes representing the horizontal and vertical directions, respectively. The coordinates are written as (x, y) [27]. The x-horizontal and y-vertical are represented by the first and second dimensions of the grid map, according to path length (PL) in equation (4):

$$PL = \sum_{i=0}^{n} \sqrt{(y_{i+1} - y_i)^2 + (x_{i+1} + x_i)^2} \qquad (4)$$

Where i =0,1,2,...,n, when i =0,the mobile robot is at the starting position s=(x_0, y_0),when i=n, the mobile robot is at the target position T= (x_n, y_n), (x_i, y_i) represents the coordinates of the current state of the mobile robot ;(x_{i+1}, y_{i+1}) represents the coordinates of the next state of the mobile robot.

The mobile robot was taught by giving it the highest reward for the inclined road from among the options available to it so that it can train on taking the inclined road if it is free of obstacles. According to the Q-learning method of learning, which relies on the principle of repetition and avoiding mistakes. The principle of reward and punishment that Q-learning relies on in reinforced learning also has a role in accelerating the learning process in the mobile robot by making the reward variable, of course, due to the change in the motor state and the presence of obstacles around the mobile robot. Whenever the angle of movement is less than 90 and the road is free of obstacles in this direction, we want the mobile robot to notice the behavior of this direction by giving it the highest reward. Through experience, Is founded the proposed method achieves a shorter time and better behavior for the mobile. It is simple to integrate into grid maps and gives mobile robots ahead of time environmental knowledge, accelerating convergence and reducing down on computation time.

3.4 Reward function

Reward and punishment are two important factors that affect how well reinforcement learning technique's function. In the Q-learning approach, the agent moves toward the next state and receives reinforcement from its environment [4]. Maximizing the total expected rewards for a sequence of states that precede the desired state is the agent's objective. Consequently, the design of the reward function has a direct bearing on how successful learning is [29]. This paper creates the reward function as equation (3) of Q-values using 8-connection.

$$R = \begin{array}{ll} 1 & Single\ Her.or\ Ver. \\ 1.4 & Single\ Diagonal \\ 3 & s_{t+1}\ is\ the\ start\ node \\ 10 & s_{t+1}\ target\ node \\ -5 & is\ the\ eight\ grids\ around\ s_{t+1}\ forbidden\ node \end{array} \quad (5)$$

4 Quick and Reassured condition of path planning

The essential concepts of the suggested Q8CQL method, which is used in this research to solve the mobile robot path planning problem, are as follows: In order to initialize the Q-table and give the mobile robot its past knowledge of the surroundings, the 8-connection with the sloped path is utilized [30]. In summary, the mobile robot uses the Q8CQL algorithm to control its distance from obstacles and travel toward the target location by combining static and dynamic rewards. The flow of the mobile robot is as follows: Choose the mode (short or secure condition path planning) once the mobile robot has reached its starting position. Using the secure condition, mobile robots are driven as far away from shoreline objects as feasible to prevent collisions with reefs. Quick condition path planning is used to find the shortest path length for mobile robots when distance and energy consumption are taken into consideration. Better path planning for mobile robots based on real-world scenarios is made possible by the integration of the two models.

5 Experimental Results and Discussion

Referring to the research [26], To create a simulation environment, we built a virtual experiment environment based on the previously mentioned maze setting. Open path, start, and goal states are examples of obstacles in a maze scenario. Figure (3) illustrates the presence of multiple barriers in a 12 x 12 grid pattern. It was simulated that a mobile robot was traveling from the starting point to the destination. The experiment aimed to find the optimal path to the destination by avoiding obstacles at the initial location where a mobile robot was defined. never veering off course or trying to figure out how to get from one point to another in order to accomplish an objective. Based on the value of the action, move to the right, left, up, up left, up right, down, down left and down right for each terminal condition as you approach obstacles. To train the robot and show that the new method works, a lot of simulation trials were carried out. This simulation was carried out with MATLAB software. Different configurations were implemented for different conditions, and the results of these different states were used to assess the effectiveness of the proposed Q8CQL-learning technique.

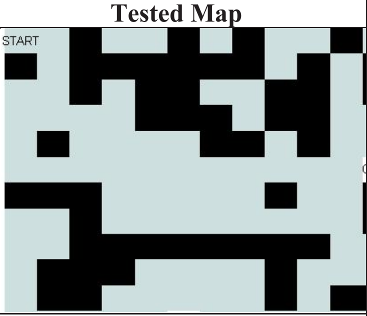
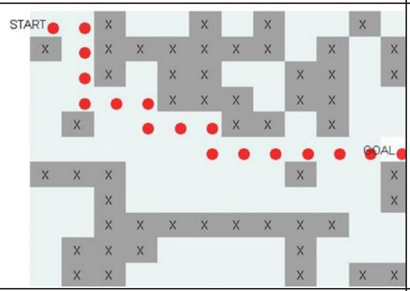
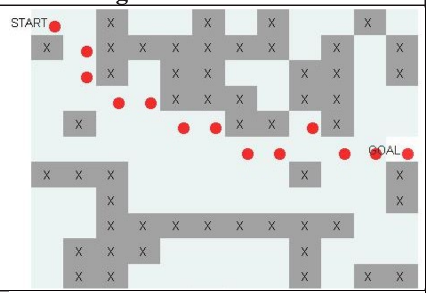
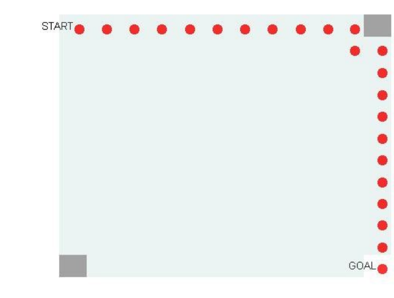
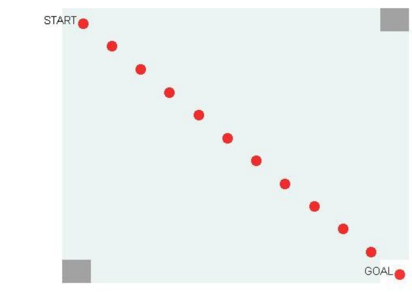
The result show in table (1) and table (2) of mobile robot path planning based on Q -learning with 4- connection and comparing the result with 8-connection. the different environment with different position of start point and goal point are use also with different angle position between them.

Table 1. The result of Q -learning with 4- connection and 8-connectionin the different environment.

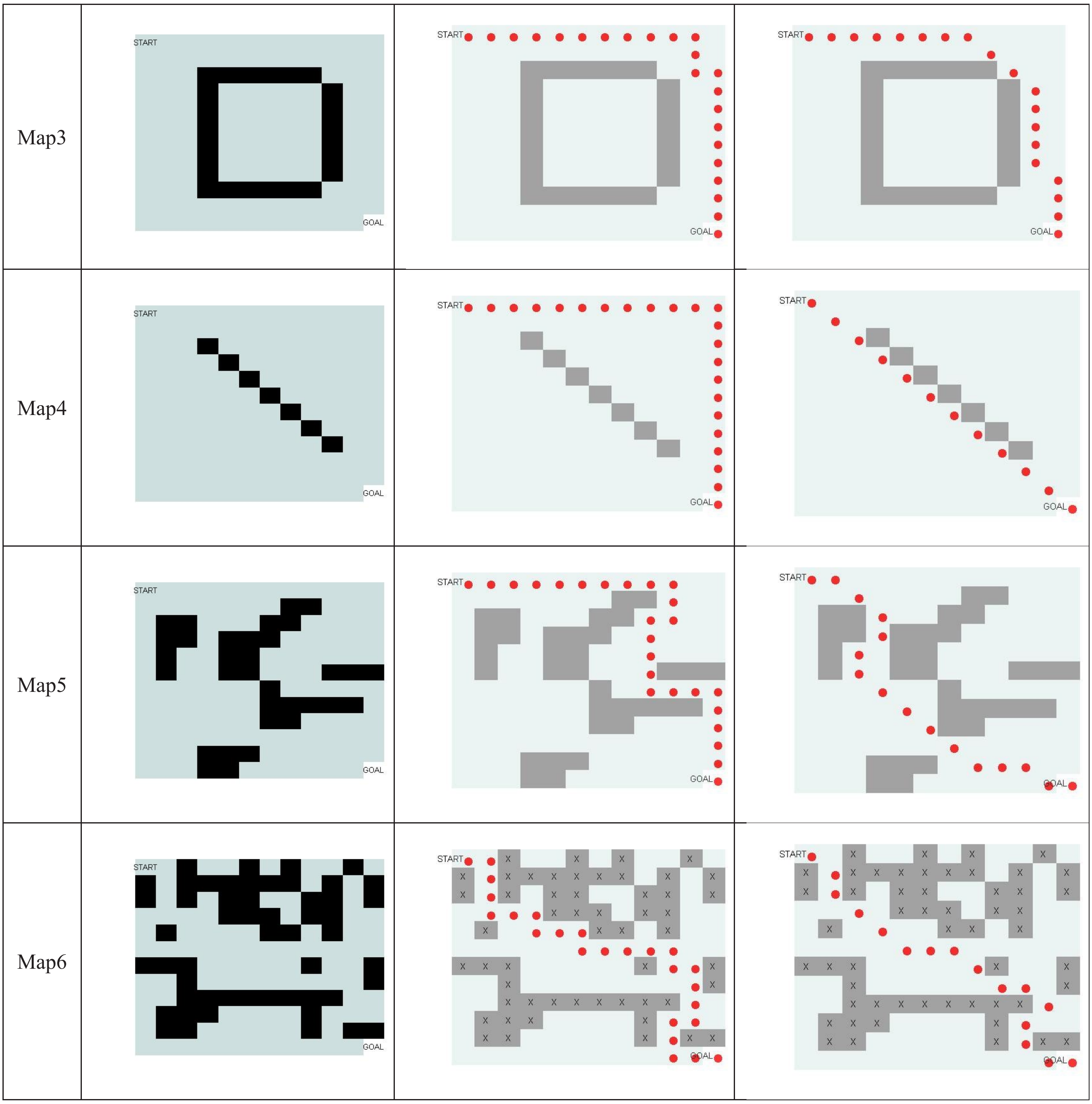
10	Direction	Mean	Distance	Total Steps	Path	Time (s)

Map1	Four	337.6863	16	17	(1 2 14 26 38 39 40 52 53 54 66 67 68 69 70 71 72)	2.116621
	Eight	446.8431	14.4	13	(1 14 26 39 40 53 54 67 68 57 70 71 72)	1.624069
Map2	Four	755.8039	22	23	(1 2 3 4 5 6 7 8 9 10 11 23 24 36 48 60 72 84 96 108 120 132 144)	1.987422
	Eight	793.6471	15.4	12	(1 14 27 40 53 66 79 92 105 118 131 144)	1.968942
Map3	Four	743.9216	22	23	(1 2 3 4 5 6 7 8 9 10 11 23 35 36 48 60 72 84 96 108 120 132 144)	1.980413
	Eight	558.098	19.6	19	(1 2 3 4 5 6 7 8 21 34 47 59 71 83 95 108 120 132 144)	1.803333
Map4	Four	929.1373	22	23	(1 2 3 4 5 6 7 8 9 10 11 12 24 36 48 60 72 84 96 108 120 132 144)	1.978153
	Eight	656.0588	15.4	12	(1 14 27 40 53 66 79 92 105 118 131 144)	1.957587
Map5	Four	964.5882	24	25	(1 2 3 4 5 6 7 8 9 10 22 34 33 45 57 69 81 82 83 84 96 108 120 132 144)	2.252676
	Eight	634.3922	18.6	16	(1 2 15 28 40 51 64 76 89 102 115 116 117 130 131 144)	1.807293
Map6	Four	1789.6078	24	25	(1 2 14 26 38 39 40 52 53 54 66 67 68 69 70 82 83 95 107 119 118 130 142 143 144)	2.950709
	Eight	1040.2353	18.6	16	(1 14 26 39 52 65 66 67 80 93 94 107 118 130 143 144)	2.443182

Table 2. Show the path planning of map1,map2,map3,map4,map5 and map6

Map	Tested Map	Four Direction Walk	Eight Direction Walk
Map1			
Map2			

 https://doi.org/10.1051/bioconf/20249700011



The result of map2 that environment not have free of obstacle, a time consumption of Q8CQL algorithm of 4-connection is 1.987 sec. and the length of path is 22 of total step of mobile robot is 23, while a time consumption of Q8CQL algorithm of 8-connection is smaller equal to 1.968 sec. and length of path is 15.4 and number of total steps of mobile robot is 12.

The result of map4 that environment have 7 obstacles, a time consumption of Q8CQL algorithm of 4-connection is 1.978 sec. and the length of path is 22 of total step of mobile robot is 2. while a time consumption of Q8CQL algorithm of 8-connection is smaller equal to 1.957 sec. and length of path is 15.4 and number of total steps of mobile robot is 12.

When increase the number of obstacles such as in environment map6 a time consumption of Q8CQL algorithm is increasing but remain of 8-connection algorithm less than 4- connection algorithm, of 4-connection is 2.950 sec. and the length of path is 24 of total step of mobile robot is 24, while a time consumption of Q8CQL algorithm of 8-connection is smaller equal to 2.443 sec. and length of path is 18.6 and number of total steps of mobile robot is 16.

The development that occurred in training the robot to choose the inclined route by giving it a smaller reward so that the robot would prefer to choose the open inclined route. This development relates to the method of searching and filling out the

Q table. However, if it continues in the above method and the route is safe along the specified path. This work will demonstrate the use of reinforcement learning as a strategy for path optimization in a maze environment and offer experimental verification of the reduction of passive exploration through the use of a simple technique to improve path finding accuracy and conserve path finding time. We constructed a simulation environment to assess path navigation in an intricate maze setting and develop Q-Learning algorithms for reinforcement learning. Subsequently, we conducted a comparison between the trial outcomes concerning the ultimate path and the path-finding durations. Now was the time to find the route to the Q8CQL algorithm faster than Q-Learning algorithm. Comparing the end paths of the two algorithms to the average results of multiple experiments, the optimization learning algorithm's path was shorter and smoother than the Q-learning algorithms.

6 Conclusion

The path searching time and path length that the Q-Learning optimization and Q-algorithm tests produced. Several tests have been conducted to try to improve the results using the Q8CQL-Learning method, but we have found that it is not very good at finding search paths in a complicated environment with plenty of obstacles. One of the characteristics of the Q8CQL Leaning algorithm is its capacity to determine the path at random. minimizing this random experience and determining the quickest and safest path. The agent's performance and efficiency were enhanced by the higher dynamic reward of the safe path of optimization-Q-learning during action, which was determined by the movement's target position and the high reward value surrounding it. Agents typically look into regions with high yield values.

It could take a while to learn anything new. Since the inclined path is always shorter and gets to the destination faster due to the law of distances, the Q8CQL method was created to select the safest and fastest path by training the robot during the education phase and integrating the decision-making phase for the next state. The inclined path receives the highest reward in each case. This fundamental research serves as an example of how to enhance the surrounding area, target mobility, or obstructions. To enable the implementation of this technology in a dynamic context, a new concept of state space has been devised.

Reference

1. P. Wang, Ch. Chan and A. de-L. Fortelle, A Reinforcement Learning Based Approach for Automated Lane Change Maneuvers, 2018 IEEE Intelligent Vehicles Symposium (IV), China, 2018.
2. M. Naeem, S.T.H. Rizvi, and A. Coronato, A Gentle Introduction to Reinforcement Learning and Its Application in Different Fields", IEEE ACCESS 8, 2020.
3. G.U.O. Tong, N. Jiang, L.I. Biyue, Z.H.U. Xi, Y. Wang, UAV navigation in high dynamic environments: A deep reinforcement learning approach, Chinese Journal of Aeronautics, Production and hosting by Elsevier 2020.
4. Parag Kulkarni, "Reinforcement and Systemic Machine Learning for Decision Making", Published by John Wiley & Sons, Inc., Hoboken, Published simultaneously in Canada (IEEE series on systems science and engineering; ISBN 978-0-470-91999-6.
5. chris Watkins, "Learning from Delayed Rewards", thesis submitted for phd in, king's college, London,1989.
6. Bakr S. Shihab, Hadeel N. Abdullah and Layth A. Hassnawi, 'Obstacle Avoidance and Path Planning for UAV Using Laguerre Polynomial', International Journal of Intelligent Engineering and Systems, Vol.15, No.6, 2022, DOI: 10.22266/ijies2022.1231.58
7. Marco Wiering and Martijn van Otterlo (Eds.),' Reinforcement Learning State-of-the-Art', ISSN 1867-4534, ISBN 978-3-642-27644-6, DOI 10.1007/978-3-642-27645-3, Springer
8. F.A. Raheem, A.T. Sadiq, N. A. F. Abbas, Robot Arm Free Cartesian Space Analysis for Heuristic Path Planning Enhancement", International Journal of Mechanical & Mechatronics Engineering IJMME-IJENS 19(01) 2019.
9. A.T. Sadiq, and A. H. Hasan, Robot Path Planning Based on PSO and D* Algorithms in Dynamic Environment, International Conference on Current Research in Computer Science and Information Technology (ICCIT), Slemani – Iraq 2017
10. Bakr S. Shihab, Hadeel N. Abdullah and Layth A. Hassnawi, 'Improved Artificial Bee Colony Algorithm-based Path Planning of Unmanned Aerial Vehicle Using Late Acceptance Hill Climbing', International Journal of Intelligent Engineering and Systems, Vol.15, No.6, 2022, DOI: 10.22266/ijies2022.1231.39 .
11. Nuha H. Abdulghafoor and Hadeel N. Abdullah, 'A novel real-time multiple objects detection and tracking framework for different challenges', Alexandria Engineering Journal, Volume 61, Issue 12, December 2022, Pages 9637-9647.

12. Nuha H. Abdulghafoor and Hadeel N. Abdullah, ‘Multiple Object Detection, Multiple Object Tracking, Classification, Deep-learning, Principal Component Pursuit’, Alexandria Engineering Journal, 2020.
13. J. Qin, X. Han, and G. Liu et al,” Path Planning Method of Mobile Robot Based on Q-learning, Journal of Physics: Conference Series,2022.
14. H.S. Lee, and J. Jeong, Mobile Robot Path Optimization Technique Based on Reinforcement Learning Algorithm in Warehouse Environment, Appl. Sci.11(2021)1209. https://doi.org/10.3390/ app11031209
15. R.S. Sutton, and A.G. Barto, Introduction to Reinforcement Learning, 2nd ed.; MIT Press: London, UK (2018) 1–528
16. B. Jang,M. Kim, G. Harerimana, J.W. Kim, Q-Learning Algorithms: A Comprehensive Classification and Applications. IEEE 7 (2019) 133653–133667
17. W.-G. Han, S.M. Baek, and T.Y. Kuc, Genetic algorithm-based path planning and dynamic obstacle avoidance of mobile robots. In Proceedings of the 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, Orlando, FL, USA 1997.
18. Siding Li, Xin Xu, Lei Zuo, ‘Dynamic Path Planning of a Mobile Robot with Improved Q-Learning algorithm’, Proceeding of the 2015 IEEEInternational Conference on Information and Automation, Lijiang, China, August 2015.
19. Ee Soong Low, Pauline Ong, Kah Chun Cheah, Solving the optimal path planning of a mobile robot using improved Q-learning’, Robotics and utonomous Systems journal homepage: www.elsevier.com/locate/robot , in 2019.
20. Monira essa aloud and nora al khamees, ,’ Intelligent Algorithmic Trading Strategy Using Reinforcement Learning and Directional Change’ ,IEEE ACCESS, 2021
21. Talal Bonny and Mariam Kashkash, ‘Highly optimized Q-learning-based bees’ approach for mobile robot path planning in static and dynamic environments’, J Field Robotics. 2022; 39:317–334.4, wileyonlinelibrary.com/journal/rob, 2022.
22. He Du, Bing Hao,Jian shuo Zhao, Jiamin Zhang, Qi Wang, Qi Yuan, ‘Apathplanning approach for mobile robots using short and safe Q-learning’, PLOSONE, https://doi.org/10.1371/journal.pone.0275100, September 26, 2022.
23. N. Kohl, and P. Stone, Policy gradient reinforcement learning for fast quadrupedal locomotion. In: International Conference on Robotics and Automation. IEEE 2004.
24. Nadia I. Khalil, Hadeel N. Abdullah and Layth A. Hassnawi, ‘A Hybrid Modified ABC-PSO Algorithm for Optimal Robotic Path Planner’, Date Added to IEEE Xplore: 17 April 2023, DOI: 10.1109/DeSE58274.2023.10100021.
25. M.A.K. Jaradat, M. Al-Rousan, L. Quadan, Reinforcement based mobile robot navigation in dynamic environment”, Robotics and Computer-Integrated Manufacturing 27 (2011).
26. M. Kirtas, K. Tsampazis, N. Passalis, A. Deepbots, A Webots-Based Deep Reinforcement Learning Framework for Robotics. In Proceedings of the 16th IFIP WG 12.5 International Conference AIAI 2020, Marmaras, Greece, (2020) 64–75
27. V.N. Sichkar, Reinforcement Learning Algorithms in Global Path Planning for Mobile Robot. In Proceedings of the 2019 International Conference on Industrial Engineering Applications and Manufacturing, Sochi, Russia 2019.
28. F.A. Raheem, A.T. Sadiq, N. A. F. Abbas, Optimal Trajectory Planning of 2-DOF Robot Arm Using the Integration of PSO Based on D* Algorithm and Cubic Polynomial Equation”, the first for Conference engineering researches 2017.
29. L. Jiang, R. Wei and D. Wang, Multi-UAV Roundup Inspired by Hierarchical Cognition Consistency Learning Based on an Interaction Mechanism, Drones 7(2023) 462.
30. S.S. Mousavi, M. Schukat, and E. Howley, Traffic light control using deep policy gradient and value-function-based reinforcement learning, IET The Institution of Engineering and Technology 2017.