ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

UDC 004.896 DOI: https://doi.org/10.30837/ITSSI.2020.13.043

V. MARTOVYTSKYI, O. IVANIUK

APPROACH TO BUILDING A GLOBAL MOBILE AGENT WAY BASED ON Q-LEARNING

Today, the problem of navigation of autonomous mobile systems in a space where disturbances are possible is urgent. The task of finding a route for a mobile robot is a complex and non-trivial task. At the moment, there are many algorithms that allow you to solve such problems in accordance with the specified criteria for building a route. Most of these algorithms are modifications of "basic" path planning methods that are optimized for specific conditions. The **subject** of research in the article is the process of building a global path for a mobile agent. The **purpose** of the work is to create an algorithm for planning the route of autonomous mobile systems in space using the Q-learning algorithm. The following **tasks** are solved in the article: development of an approach to training and support of a reinforcement learning algorithm for building a global path of a mobile agent; testing the agent's ability to find a path in environments that are not in the training set. The following **methods** are used: graph theory, queuing theory, Markov decision-making process theory and mathematical programming methods. The research is based on scientific articles and other materials from foreign conferences and archives in the field of machine learning, deep learning and deep reinforcement learning. The following **results** were obtained: an approach was formulated to construct the global path of a mobile agent based on the accumulated data in the process of interaction with the external environment. The environment rewards these actions and the agent continues to carry them out. This approach will allow this method to be applied to a wide range of situations and devices. **Conclusions**: This approach allows accumulating the knowledge of the outside world for further decision-making when planning a route where the robot can acquire the skill of self-learning, studying and training like a human, and finding the path from the initial state to the target state in an unknown environment. In the modern world, the use of robots and autonomous systems is spreading, designed to replace or facilitate human labor, make it safer and speed it up. Adaptive autonomous path finding algorithms are very important in many robotics applications. Thus, navigation tasks with limited information are relevant today, since this is the main task that the agent solves, and one of the tasks that are part of the robot during operation.

Keywords: path planning; Q-learning; mobile works; adaptive standalone search algorithms.

Introduction

Today, the problem of navigation of autonomous mobile systems in a space where disturbances are possible is urgent. The problem lies in the fact that various disturbances arising during the movement of work do not allow the implementation of movement along a pre-planned route and require current redevelopment in accordance with the situation received from the sensors. For autonomous systems, the problem is aggravated by the need to automatically generate a model of the current situation based on data from sensors and to integrate this model of the situation with planning and control models in real time [1].

In this work, by planning and control problems we mean the problem of finding the optimal sequence of actions, which leads to the agent getting from the initial position to the final one. At the same time, at every step the agent receives information about the environment. This information may be complete or incomplete.

Complete information is that information that fully describes the state of the agent together with the environment [2]. In planning and control tasks, this can be a map with a marked position of the agent on it. To solve the navigation problem with complete information, one of the classical search algorithms, such as Dijkstra's A* algorithm, and their modifications, and the like, can almost always be applied. However, in real conditions such a map is very difficult or impossible to build, and therefore most often we have only incomplete information. Incomplete information in search tasks is usually data from sensors at work or from some static structures in the environment.

Effective mobile operations in 3-dimensional space are the important research topic in artificial intelligence. In the modern world, the use of robots on autonomous systems is spreading, designed to replace or facilitate human labor, make it safer and faster. Adaptive autonomous pathfinding algorithms are very important in many applications of robotics.

For example, it is very important for security workers in the fire, rescue and police services to ensure their own safety while performing tasks. They independently penetrate into dangerous environments: apartments, houses, premises of various types. Special works of varying degrees of autonomy are disposable for a long time in developed countries [3].

Many people use home helpers today. Robot vacuum cleaners, voice assistants, smart home systems are getting smarter every day, demonstrating advances in understanding human speech, navigating the home, monitoring room performance, and the like. From the navigation side, most of them are arranged quite simply and rely on classical algorithms for building maps and planning using them.

Analysis of the problem and existing methods

The task of finding a route for a mobile robot is a complex and non-trivial task. At the moment, there are many algorithms that allow you to solve such problems in accordance with the specified criteria for building a route. Most of these algorithms are modifications of "basic" path planning methods that are optimized for specific conditions.

Pathfinding algorithms can be divided into 3 unique groups:

1. graph-based algorithms;

2. algorithms for avoiding obstacles;
3. algorithms using intelligent methods.

Analysis of the literature has shown that recently, many different methods and algorithms have been proposed for route planning [4–14]. The article [6] presents a solution for planning the shortest route for moving a robot in a maze based on the Voronoi graph and has the following form in fig. 1. Algorithm represented in fig. 1 is based on the use of the method and successfully solves the problem of finding the optimal route using a set of optimality criteria. The representation method is based on the Voronoi graph and helps to avoid the problem of path getting stuck during iteration at local minima and provides more flexibility for route optimization.

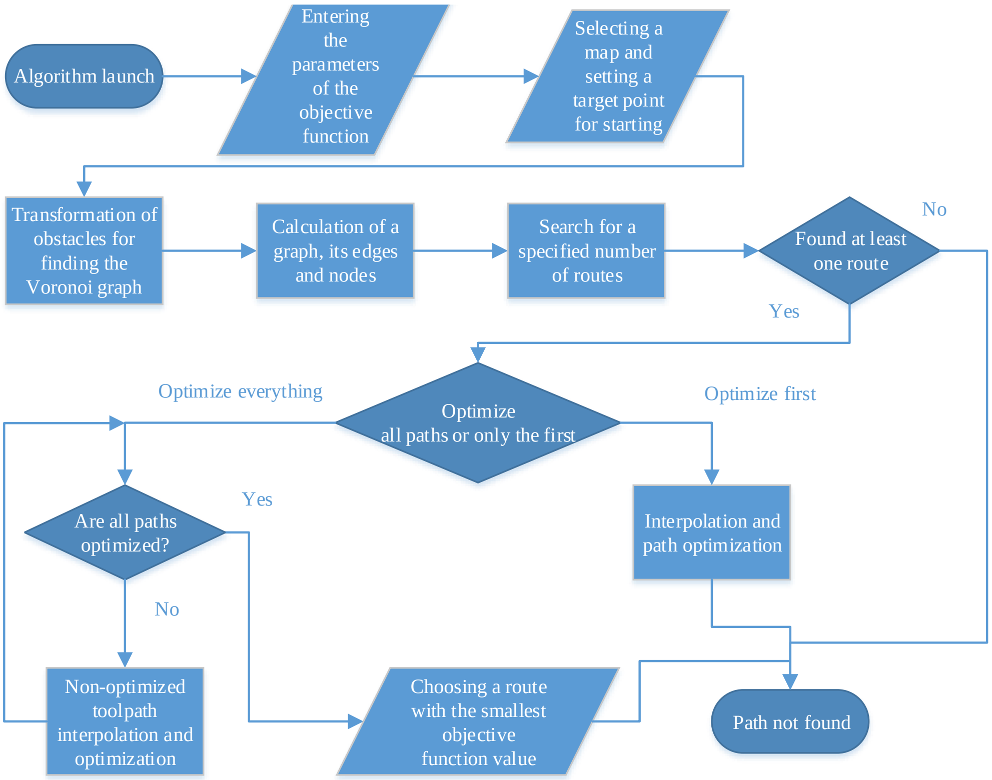
Algorithm launch
Entering the parameters of the objective function
Selecting a map and setting a target point for starting
Transformation of obstacles for finding the Voronoi graph
Calculation of a graph, its edges and nodes
Search for a specified number of routes
Found at least one route
No
Yes
Optimize everything
Optimize all paths or only the first
Optimize first
Are all paths optimized?
Yes
Interpolation and path optimization
No
Non-optimized toolpath interpolation and optimization
Choosing a route with the smallest objective function value
Path not found

Fig. 1. Scheme of the method for finding the optimal route based on the Voronoi graph

In [7], a global approach to route planning is implemented using artificial potential fields for multi-robot systems (MRS). A 3D lead map is created using simplified lead functions. To create a three-dimensional map, both the gravitational forces between the robots and the target and the repulsive forces that push the robots away from obstacles and each other are calculated. The problem of local minima is solved by using a virtual obstacle approach. The path of the robot is formed, starting from the initial position of the robot to the target, based on the generated 3D potential map that the mobile robots should follow.

The article [9] proposes a method based on a genetic algorithm for planning articulated-type mobile robots. The proposed algorithm considers path planning as a multipurpose optimization problem and evaluates the effectiveness of the result based on four adjustable fitness objective functions. The algorithm generates optimal ones according to the Pareto principle. Fig. 2 shows a block diagram of the proposed method based on a genetic algorithm. The method takes as input an obstacle map (W), a roadmap (Q) and several parameters associated with the genetic algorithm, which gives the ideal sequence of commands for the movement of the robot. The roadmap (Q) is a series of predefined configurations of the hTetro (q) robot, which defines a series of positions and morphologies that the robot should arrive at during the navigation process.

In [10], an algorithm for optimizing ant colonies based on optimizing a swarm of particles is used to find the optimal route. Due to various limitations such as limited battery power and limited visibility, the ant colony algorithm uses an improved pheromone update rule and a heuristic function based on the particle swarm optimization algorithm. The solution to the route planning process is described as follows:

Step 1. The starting point and the target point in the abstract model of the environment are determined first after building a three-dimensional model of the environment and determining the main direction of movement of the ant.

Step 2. Based on the heuristic information and the weight of the pheromone value, the next search point for ants is determined by formula (10) in article [10].

Step 3. Then the local pheromone footprint is updated in accordance with the formula (11) [10]

Step 4. Determine if all the ants have completed the construction of the trail. If not, then go back to Step 2.

Step 5. The global pheromone footprint is updated according to equation (14) in [10] to determine if the algorithm satisfies the stop condition. Otherwise, go back to Step 2.

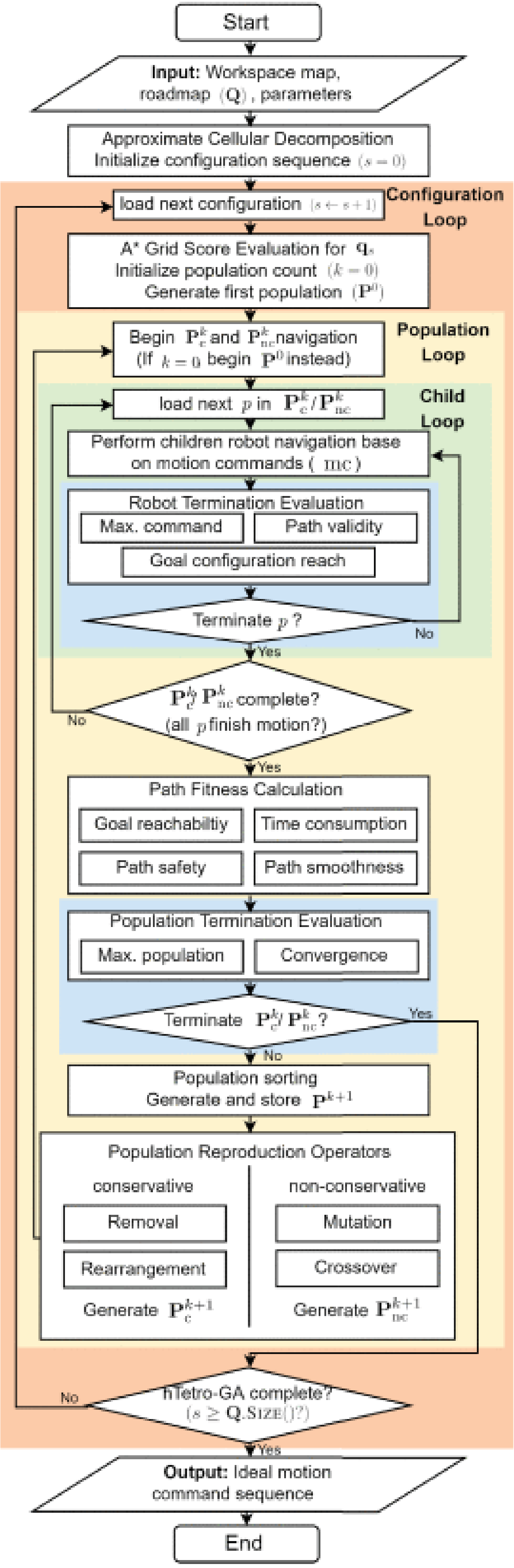
Start
Input: Workspace map, roadmap (Q), parameters
Approximate Cellular Decomposition
Initialize configuration sequence (s = 0)
Configuration Loop
load next configuration (s ← s + 1)
A* Grid Score Evaluation for q_s
Initialize population count (k = 0)
Generate first population (P^0)
Population Loop
Begin P_c^k and P_{nc}^k navigation
(If k = 0 begin P^0 instead)
Child Loop
load next p in P_c^k/P_{nc}^k
Perform children robot navigation base on motion commands (mc)
Robot Termination Evaluation
Max. command
Path validity
Goal configuration reach
Terminate p ?
No
Yes
P_c^k/ P_{nc}^k complete?
(all p finish motion?)
No
Yes
Path Fitness Calculation
Goal reachabiltiy
Time consumption
Path safety
Path smoothness
Population Termination Evaluation
Max. population
Convergence
Terminate P_c^k/ P_{nc}^k ?
Yes
No
Population sorting
Generate and store P^{k+1}
Population Reproduction Operators
conservative
non-conservative
Removal
Mutation
Rearrangement
Crossover
Generate P_c^{k+1}
Generate P_{nc}^{k+1}
hTetro-GA complete?
(s ≥ Q.SIZE()?)
No
Yes
Output: Ideal motion command sequence
End

Fig. 2. Scheme of the hTetro-GA algorithm [9]

Analysis of existing solutions reflects the individuality of using the developed methods for a specific situation and device.

Formulation of the problem

This paper proposes an approach to building a global path for a mobile agent based on the accumulated data in the process of interacting with the external environment. The environment rewards these actions and the agent continues to carry them out. This approach will allow this method to be applied to a wide range of situations and devices.

The approach is as follows: there is a mobile agent that interacts with the external environment described in the form of a Markov decision making process (MDMP) takes one of a predetermined set of actions. Using partial learning algorithms, we are trying to find a strategy that assigns actions to the states of the environment, one of which the agent can choose in these states and achieve the maximum reward. Interaction with the environment is shown in fig. 3.

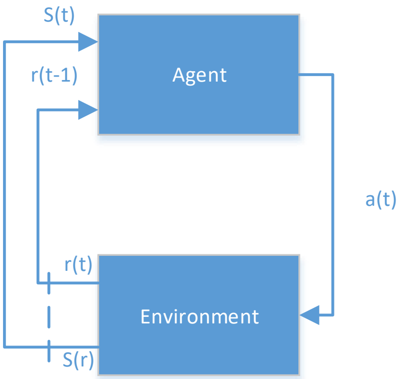

Fig. 3. Interaction of the agent with the environment

Formally, the approach to building a global mobile agent path based on Q-learning is described as follows:

- the set of states of the external environment S;
- a set of actions A;
- lots of scalar "rewards".

At an arbitrary time instant *t*, the agent is characterized by a state $s_t \in S$ and a set of all possible actions in the current state of the environment $A(s_t)$. Making a choice of action $a \in A(s_t)$, the agent goes into a state s_{t+1} and receives a payoff r_t. Based on such interaction with the environment, the agent, thanks to Q-learning, must develop a strategy $\Omega : S \to A$ that maximizes the amount of reward $R = r_0 + r_1 + \cdots + r_n$ in the case of an MDMP having a terminal state, or the value:

$$R = \sum_t \gamma^t r_t \,, \tag{1}$$

for MDMP without terminal states (where $0 \le \gamma \le 1$ is the discount factor for "expected reward").

The general learning algorithm is shown in fig. 4, where

$$P(s_{t+1} = s', r_{t+1} = r \mid s_t, a_t, r_t, s_{t-1}, a_{t-1}, r_{t-1}, .., s_1, a_1) = \\ = P(s_{t+1} = s', r_{t+1} = r \mid s_t, a_t) \,. \tag{2}$$

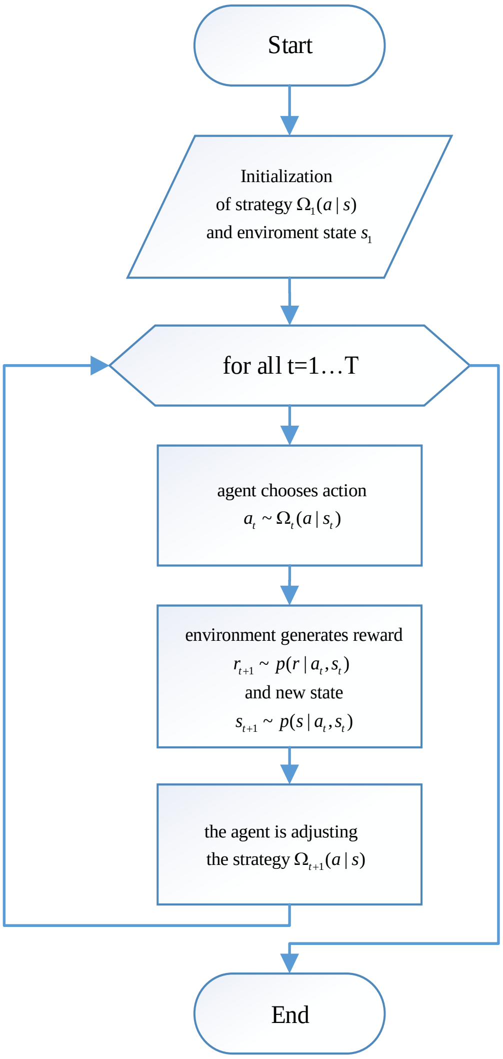
Start
Initialization
of strategy $\Omega_1(a \mid s)$
and enviroment state s_1
for all t=1…T
agent chooses action
$a_t \sim \Omega_t(a \mid s_t)$
environment generates reward
$r_{t+1} \sim p(r \mid a_t, s_t)$
and new state
$s_{t+1} \sim p(s \mid a_t, s_t)$
the agent is adjusting
the strategy $\Omega_{t+1}(a \mid s)$
End

Fig. 4. Learning algorithm

Building a global mobile agent path based on *Q*-learning.

Based on the reward that the agent receives from the external environment, the utility function *Q* is formed, which subsequently makes it possible not to randomly choose a strategy of behavior, but to take into account the accumulated experience of previous interaction with the external environment.

Thus, the algorithm is a function of quality from state and action:

$$Q : S \times A \to \mathbb{R} \,. \tag{3}$$

Before training, *Q* is initialized with random values. After that, at each moment of time t, the agent chooses an action a_t, receives a reward r_t, switches to a new state s_{t+1}, which may depend on the previous state s_t and the selected action, and updates the function *Q*. Updating the function uses a weighted average between the old and new values:

$$Q^{new}(s_{t+1},a_{t+1}) = Q(s_t,a_t) + \\ +\lambda*(r_t+\gamma*\max_a(Q^{new}(s_{t+1},a_{t+1})-Q(s_t,a_t))), \qquad (4)$$

where $T = S\times A \to S$ is a transition function, $Q^{new}(s_{t+1},a_{t+1})$ is the value of the objective function in the next step, $Q(s_t,a_t)$ – the value of the objective function at the current position, $\max_a(Q^{new}(s_{t+1},a_{t+1})-Q(s_t,a_t))$ – selection of the maximum value from the possible next steps, $s\in S$ – agent's current position, $a\in A$ – current action, $\lambda\in[0,1]$ – the speed of learning, the higher it is, the more the agent trusts new information, $R = S\times A\to\mathbb{R}$ – reward function, $r_t \in R$ – the reward received in the current position, $\gamma\in[0,1]$ – gamma (decrease in remuneration, discount factor), the smaller it is, the less the agent thinks about the benefits of his future actions, s_{t+1} – the next selected position according to the next selected action, a_{t+1} – next selected action.

The main element in the Q-learning approach is the reward matrix - the Q-table of the state of the system. Matrix Q is a set of system states and weights of the system's response to various actions. While trying to get through the given environment, the mobile robot learns to avoid obstacles and find its way to its destination. As a result of the interaction between the agent and the external environment, a Q-table of accumulated experience is built, with the help of which the mobile robot decides on the next step.

The algorithm for accumulating knowledge from the external environment is shown in fig. 5.

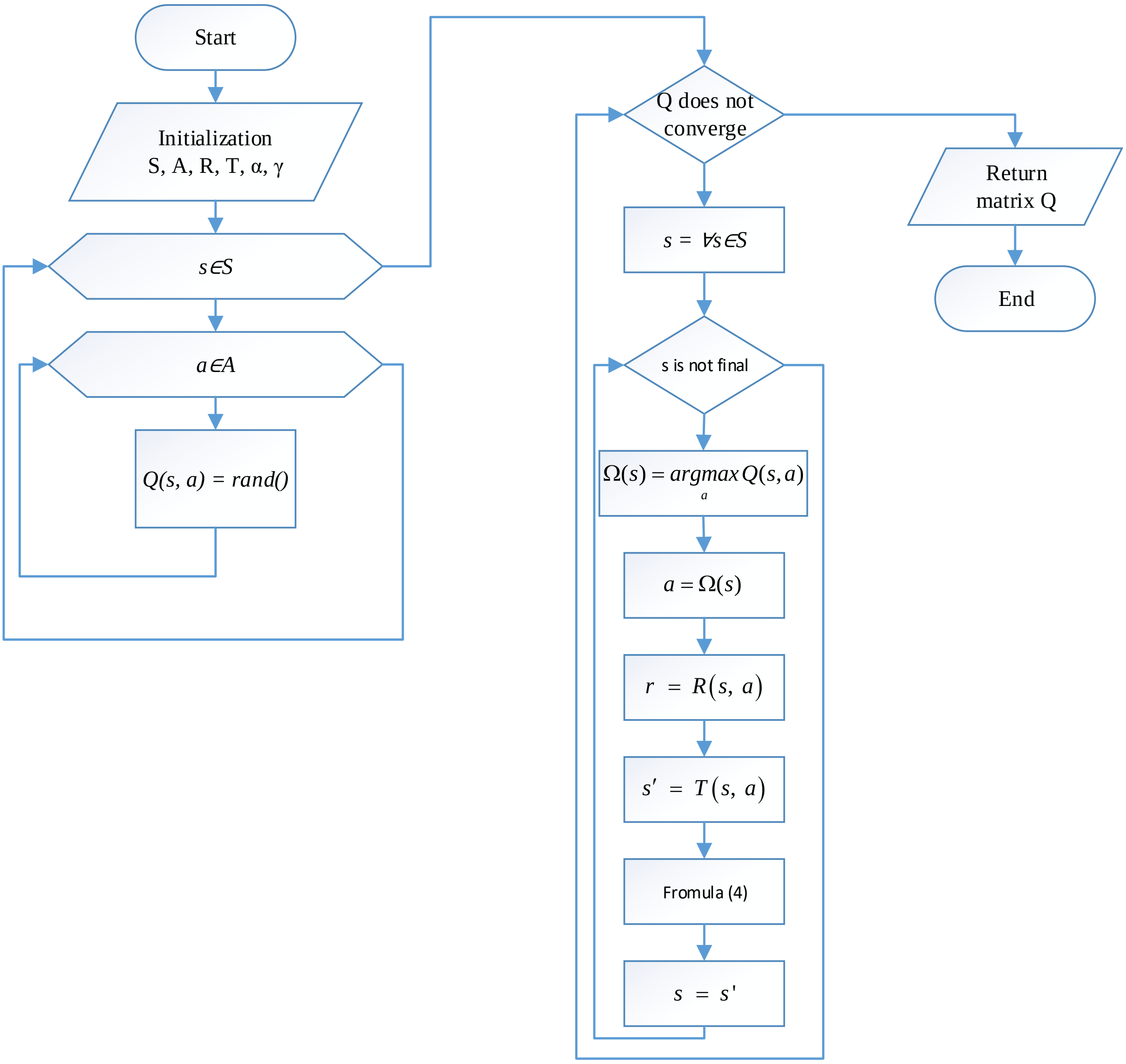

Fig. 5. Q-learning

One of the benefits of Q-learning is that it is able to compare the expected utility of available activities without shaping environmental models it is used for situations that can be represented as MDMP.

Experiment results

In order to test the approach based on Q-learning, a software implementation of the algorithm in the Python programming language was made. During the simulation, a computer with the following characteristics was used: Intel Core i5-4300U 2.5 GHz processor, 8Gb RAM.

The software implementation of the algorithm consists of two modules: the agent module and the environment module. The agent affects the external environment, and the external environment in response to the agent's actions affects him.

At each step, the agent: performs an action (up, down, right, left); receives observation (new state), receives a reward. Wednesday: receives observation (new state), issues reward.

The algorithms showed the fastest convergence with the following parameters: λ = 0.5; γ = 0.99. The result of the route planning simulation is shown in fig. 6.

For the proposed model, graphs of the main indicators during training are shown: average, minimum and maximum rewards and entropy of the resulting strategy in figs. 7, 8.

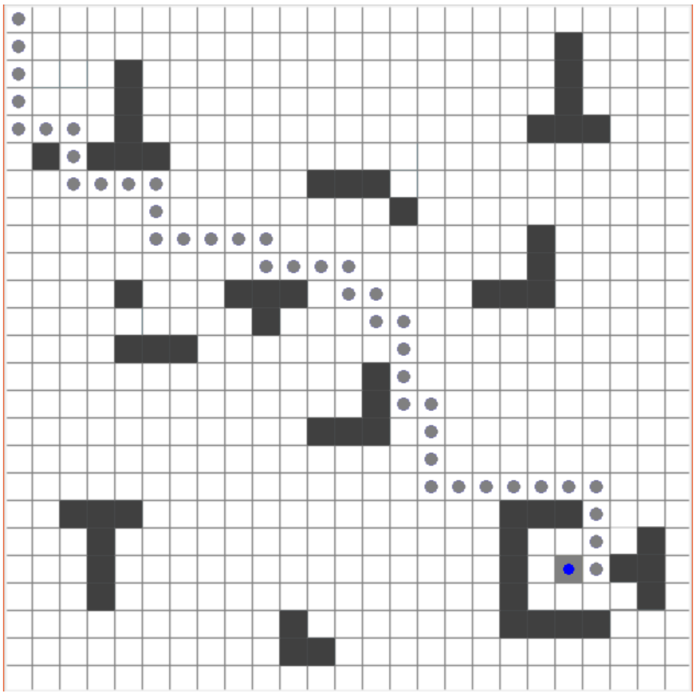

Fig. 6. Example of a constructed route

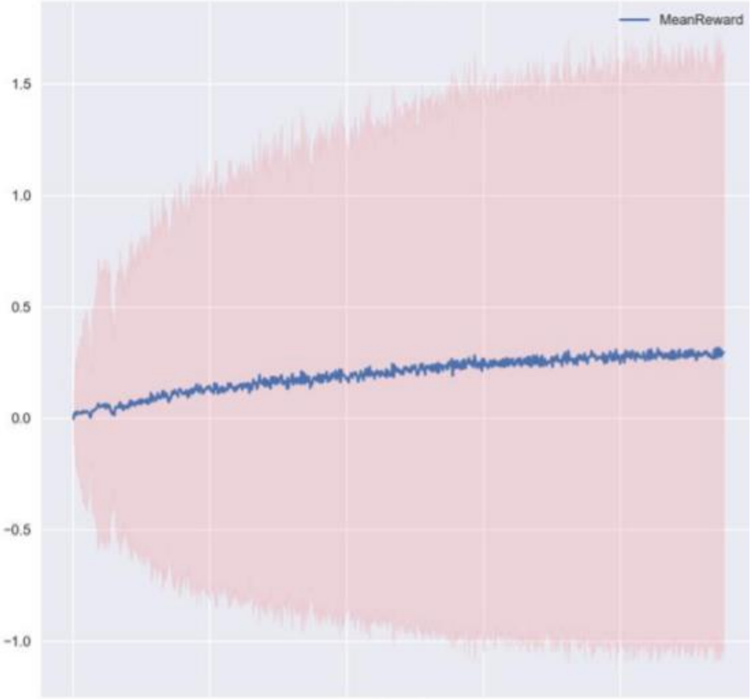

Fig. 7. Average, minimum and maximum training rewards

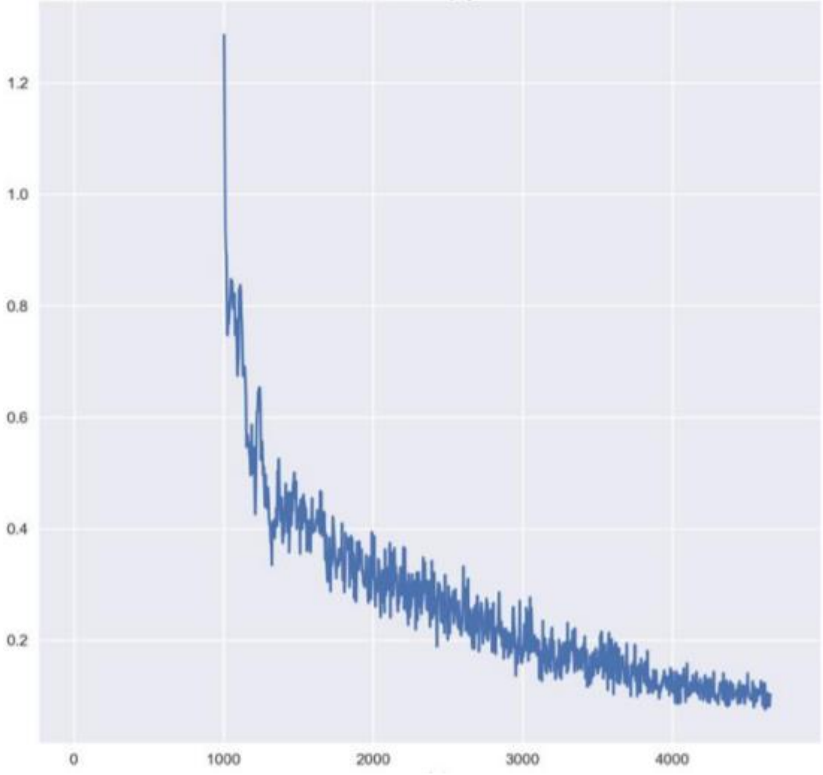

Fig. 8. Entropy of the resulting strategy

The optimal "Q-table" has values that allow the mobile robot to take the best action in each state, resulting in the best route. Values from "Q-table" are used to create a strategy when planning a route between two points. In this case, it would be a greedy strategy, because the mobile robot always takes the action it thinks is best in each state.

Conclusions

This article presents an approach to building a global mobile agent path based on Q-learning. When changing the local environment, Q-learning uses local path planning. Thus, the robot can acquire the skill of self-learning by learning and training like a human and looking for a clear path from the initial state to the target state in an unknown environment.

The work developed and trained a decision-making system associated with high rates of success, finds short paths to the specified target points in scenes with both training and variation sampling, demonstrating the ability to generalize.

References

1. Kargin, A., Ivaniuk, O. (2020), "Autonomous robot motion control situational planning model", *Advanced Information Systems*, No 4, P. 41–51. DOI: 10.20998/2522-9052.2020.3.05

2. Sukharev, O. (2019), "Functions of information and modes of informational development of control systems", *Problems of theory and practice of management*, No 1, P. 37–51.

3. Miyazawa, K. (2002), "Fire robots developed by the Tokyo Fire Department", *Advanced Robotics*, P. 553–556. DOI: 10.1163/156855302320535953

4. Magid, E., Lavrenov, R., Afanasyev, I. (2017), "Voronoi-based trajectory optimization for UGV path planning", *International Conference on Mechanical, System and Control Engineering (ICMSC)*, *IEEE*, P. 383–387. DOI: 10.1109/ICMSC.2017.7959506

5. Kovács, G., Yusupova, N., Smetanina, O., Rassadnikova, E. (2018), "Methods and algorithms to solve the vehicle routing problem with time windows and further conditions", *Pollack Periodica*, No. 13 (1), P. 65–76. DOI: 10.1556/606.2018.13.1.6

6. Lavrenov, R. O., Magid, E. A. (2020), "Multihomotopic search for the optimal route for autonomous mobile devices", *Industrial automation*, No. 7, P. 61–64. DOI: 10.25728/avtprom.2020.07.14

7. Hassan, A. M., Elias, C. M., Shehata, O. M., Morgan, E. I. (2017), "A global integrated artificial potential field/virtual obstacles path planning algorithm for multi-robot system applications", *Int. Research J. of Eng. and Technology*, No. 4 (9), P. 1198–1204.

8. Wahid, N., Zamzuri, H., Amer, N. H., Dwijotomo, A., Saruchi, S. A., Mazlan, S. A. (2020), "Vehicle collision avoidance motion planning strategy using artificial potential field with adaptive multi-speed scheduler", *IET Intelligent Transport Systems*, No. 14 (10), P. 1200–1209. DOI: 10.1049/iet-its.2020.0048

9. Ku Ping Cheng, Rajesh Elara Mohan, Nguyen Huu Khanh Nhan, Anh Vu Le (2020), "Multi-Objective Genetic Algorithm-Based Autonomous PP for Hinged-Tetro Reconfigurable Tiling Robot", *IEEE Access*, available at : https://ieeexplore.ieee.org/abstract/document/9131750 (last accessed: 23.09.2020)

10. Che Gaofeng, Lijun Liu, Zhen Yu (2020), "An improved ant colony optimization algorithm based on particle swarm optimization algorithm for path planning of autonomous underwater vehicle", *Journal of Ambient Intelligence and Humanized Computing*, No. 11 (8), P. 3349–3354. DOI: 10.1007/s12652-019-01531-8

11. Raja, P., Pugazhenthi, S. (2012), "Optimal path planning of mobile robots: A review", *International journal of physical sciences*, No. 7 (9), P. 1314–1320. DOI: 10.5897/IJPS11.1745

12. Yahja, A. (1998), "Framed-quadtree path planning for mobile robots operating in sparse environments", *In: Proceedings. IEEE International Conference on Robotics and Automation (Cat. No. 98CH36146)*, P. 650–655. DOI: 10.1109/ROBOT.1998.677046

13. Montiel, O., Orozco-Rosas, U., Sepúlveda, R. (2015), "Path planning for mobile robots using Bacterial Potential Field for avoiding static and dynamic obstacles", *Expert Systems with Applications*, No. 42 (12), P. 5177–5191. DOI: 10.1016/j.eswa.2015.02.033

14. Raja, P., Pugazhenthi, S. (2009), "Path planning for mobile robots in dynamic environments using particle swarm optimization", *2009 International Conference on Advances in Recent Technologies in Communication and Computing. IEEE*, P. 401–405. DOI: 10.1109/ARTCom.2009.24

15. Kovács, B., Szayer, G., Tajti, F., Burdelis, M., Korondi, P. (2016), "A novel potential field method for path planning of mobile robots by adapting animal motion attributes", *Robotics and Autonomous Systems*, No. 82, P. 24–34. DOI: 10.1016/j.robot.2016.04.007.

Received 14.08.2020

Відомості про авторів / Сведения об авторах / About the Authors

Мартовицький Віталій Олександрович – кандидат технічних наук, Харківський національний університет радіоелектроніки, доцент кафедри електронних обчислювальних машин, Харків, Україна; email: martovytskyi@gmail.com; ORCID: https://orcid.org/0000-0003-2349-0578.

Мартовицкий Виталий Александрович – кандидат технических наук, Харьковский национальный университет радиоэлектроники, доцент кафедры электронных вычислительных машин, Харьков, Украина.

Martovytskyi Vitalii – PhD (Engineering Sciences), Kharkov National University of Radio Electronics, Assistant professor of the Department of Electronic Computers, Kharkov, Ukraine.

Іванюк Олександр Ігорович – Український державний університет залізничного транспорту, аспірант кафедри інформаційних технологій, Харків, Україна; email: ivaniuk@kart.edu.ua; ORCID: https://orcid.org/0000-0002-4007-2215.

Иванюк Александр Игоревич – Украинская государственная академия железнодорожного транспорта, аспирант кафедры информационных технологий, Харьков, Украина.

Ivaniuk Oleksandr – Ukrainian State University of Railway Transport, PhD Student of the Department of Information Technology, Kharkiv, Ukraine.

ПІДХІД ДО ПОБУДОВИ ГЛОБАЛЬНОГО ШЛЯХУ МОБІЛЬНОГО АГЕНТА НА ОСНОВІ Q-LEARNING

На сьогодні актуальною є проблема навігації автономних мобільних систем в просторі, де можливі обурення. Завдання пошуку маршруту для мобільного робота - складне і нетривіальне завдання. На даний момент існує безліч алгоритмів, що дозволяють вирішувати подібні завдання відповідно до заданих критеріїв для побудови маршруту. Велика частина цих алгоритмів є модифікаціями "базових" методів планування шляху, які оптимізовані під конкретні умови. **Предметом** дослідження в статті є процес побудови глобального шляху мобільного агента. **Мета** роботи – створення алгоритму планування маршруту автономних мобільних систем в просторі з використанням алгоритму Q-learning. У статті вирішуються наступні **завдання**: розробка підходу до навчання та підтримки алгоритму навчання з підкріпленням для побудови глобального шляху мобільного агента; тестування здатності агента до пошуку шляху в середовищах, відсутніх в наборі для тренування. Використовуються такі **методи**: теорія графів, теорія масового обслуговування, теорія марковского процесу прийняття рішень і методи математичного програмування. Дослідження ґрунтується на наукових статтях і інших матеріалах зарубіжних конференцій і архівів в області машинного навчання, глибокого навчання і глибокого навчання з підкріпленням. Отримані наступні **результати**: сформульовано підхід до побудови глобального шляху мобільного агента на основі накопичених даних в процесі взаємодії із зовнішнім середовищем. Навколишнє середовище дає нагороду за ці дії, а агент продовжує їх виконувати. Такий підхід дозволить застосовувати цей метод для широкого кола ситуацій і пристроїв. **Висновки**: Даний підхід дозволяє накопичувати свої знання про навколишній світ для подальшого прийняття рішення при плануванні маршруту, де робот може отримати навик самонавчання, вивчаючись і тренуючись, як людина, та знаходити шлях від початкового стану до цільового стану в невідомому середовищі. У сучасному світі поширюється використання роботів і автономних систем, призначених замінити або полегшити людську працю, зробити її безпечнішою і прискорити її. Адаптивні автономні алгоритми пошуку шляху дуже важливі в багатьох додатках робототехніки. Таким чином, завдання навігації з обмеженою інформацією актуальні сьогодні, так як це головне завдання, яке агент вирішує, і одне із завдань, що входять до складу, виконуваних роботом при роботі.

Ключові слова: планування шляху; Q- learning; мобільні роботи; адаптивні автономні алгоритми пошуку.

ПОДХОД К ПОСТРОЕНИЮ ГЛОБАЛЬНОГО ПУТИ МОБИЛЬНОГО АГЕНТА НА ОСНОВЕ Q-LEARNING

На сегодня актуальной является проблема навигации автономных мобильных систем в пространстве, где возможны возмущения. Задача поиска маршрута для мобильного робота – сложная и нетривиальная задача. На данный момент существует множество алгоритмов, позволяющих решать подобные задачи в соответствии с заданными критериями для построения маршрута. Большая часть этих алгоритмов являются модификациями "базовых" методов планирования пути, которые оптимизированы под конкретные условия. **Предметом** исследования в статье является процесс построения глобального пути мобильного агента. **Цель** работы – создание алгоритма планирования маршрута автономных мобильных систем в пространстве с использованием алгоритма Q-learning. В статье решаются следующие **задачи**: разработка подхода к обучению и поддержке алгоритма обучения с подкреплением для построения глобального пути мобильного агента; тестирование способности агента к поиску пути в средах, отсутствующих в наборе для тренировки. Используются следующие **методы**: теория графов, теория массового обслуживания, теория марковского процесса принятия решений и методы математического программирования. Исследование основывается на научных статьях и других материалах зарубежных конференций и архивов в области машинного обучения, глубокого обучения и глубокого обучения с подкреплением. Получены следующие **результаты**: сформулирован подход к построению глобального пути мобильного агента на основе накопленных данных в процессе взаимодействия с внешней средой. Окружающая среда дает награду за эти действия, а агент продолжает их выполнять. Такой подход позволит применять этот метод для широкого круга ситуаций и устройств. **Выводы**: Данный подход позволяет накапливать свои знания о внешнем мире для дальнейшего принятия решения при планировании маршрута, где робот может получить навык самообучения, изучая и тренируясь, как человек, и нахождение путь от начального состояния к целевому состоянию в неизвестной среде. В современном мире распространяется использование роботов и автономных систем, предназначенных заменить или облегчить человеческий труд, сделать ее более безопасной и ускорить ее. Адаптивные автономные алгоритмы поиска пути очень важны во многих приложениях робототехники. Таким образом, задачи навигации с ограниченной информацией актуальны сегодня, так как это главная задача, которую агент решает, и одна из задач, входящих в состав, выполняемых роботом при работе.

Ключевые слова: планирование пути; Q-learning; мобильные роботы; адаптивные автономные алгоритмы поиска.

Бібліографічні описи / Bibliographic descriptions

Мартовицький В. О., Іванюк О. І. Підхід до побудови глобального шляху мобільного агента на основі Q-learning. *Сучасний стан наукових досліджень та технологій в промисловості*. 2020. № 3 (13). С. 43–51. DOI: https://doi.org/10.30837/ITSSI.2020.13.043.

Martovytskyi, V., Ivaniuk, O. (2020), "Approach to building a global mobile agent way based on Q-learning", *Innovative Technologies and Scientific Solutions for Industries*, No. 3 (13), P. 43–51. DOI: https://doi.org/10.30837/ITSSI.2020.13.043.