UDC 004.8

Y. Albrekht, A. Pysarenko

UNKNOWN LOCATION TARGETS SEARCHING SYSTEM IN KNOWN ENVIRONMENT USING REINFORCEMENT LEARNING

Abstract: This article investigates two different approaches to searching for objects of a certain type in a known environment: with a centrally controlled system using individual modules that transmit information and by dividing the entire search area into smaller ones and using individual objects. The article conducts experiments using reinforcement learning algorithms to compare the learning speed and capabilities of a system with search modules and centralized control and a separate object to search for static objects with random locations in a known environment and to search for objects moving at a constant speed in a known environment. The article provides detailed information about the experimental design, including the definition of the parameters for reinforcement learning and the size of the input and output data for the neural network. The results of the experiments are presented graphically, demonstrating the effectiveness of reinforcement learning and the difference in the learning speed and capabilities of the two systems under study.

Keywords: reinforcement learning, system of search modules, object detection.

Problem statement

Searching for objects of a certain type in a known environment is a fairly common task in both civilian and military life. Specific examples of such a task describe a whole range of applications: from rescuing a group of lost people and identifying trees in a garden that need to be treated for pests to detecting enemy vehicles on the battlefield. This article investigates two of the most effective approaches to conducting this search:

- with a centrally controlled system using individual modules that transmit information (e.g., a flock of drones controlled from a control center);
- by dividing the entire search area into smaller ones and using individual objects (individual drones without centralized control).

The task of object detection was reduced to the following experiments:

- using reinforcement learning algorithms, comparing the learning speed and capabilities of a system with search modules and centralized control and a separate object with a proportionally reduced area of the environment to search for static objects with random locations in a known environment;

- comparison of the learning speed and capabilities of a system with search modules and centralized control and a single object with a proportionally reduced area of the environment to search for objects moving at a constant speed in a known environment.

Analysis of recent research and publications

The basis for the experiment using reinforcement learning is research on creating algorithms for controlling objects in different environments. And in this study, the environments for the experiments were applications presented on the Atari game console [1]. For example, the experiment of reinforcement learning in the Seaquest game shown in Fig. 1. The study created a library with all environments recreated from Atari games. This library is also freely available, supported to this day, and has hundreds of different environments. Also, this library has the ability to connect various reinforcement learning algorithms and there is documentation for adding your own environments.

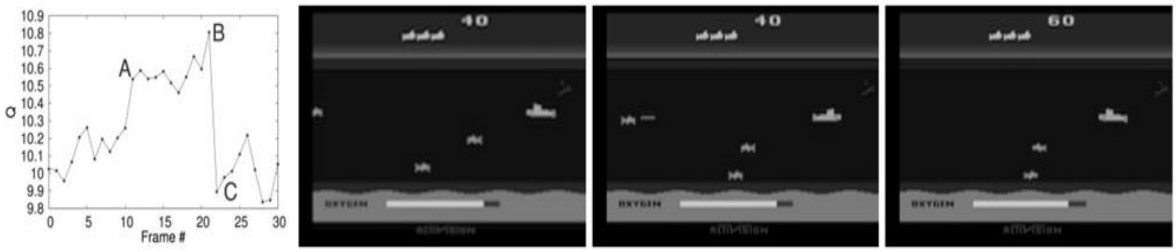

Figure 1. Demonstration of the study on the Seaquest game [1]

Study [2] defined the Swarm Dynamic Mode Decomposition algorithm, which is a modification of the Dynamic Mode Decomposition algorithm with control and allows to study the behavior of a system of modules similar to the behavior of a flock in nature. This algorithm was used in the study [3] to demonstrate the behavior of a flock when defending against an attack from another flock.

Also, possible ways to use the reinforcement learning algorithm and the problems of using a flock of drones are described in [4].

The article [5] proposes a distributed reinforcement learning (RL) approach to enable a swarm of robots to learn collective behaviors. The approach involves the use of a two-layer architecture, where the first layer coordinates the behavior of individual robots, while the second layer coordinates the behavior of the swarm. The authors tested the proposed approach in a simulation environment, where a swarm of robots was trained to move towards a target location while avoiding obstacles. The results showed that the swarm was able to learn the desired behavior and outperformed individual robots that were trained using a centralized RL approach. The authors suggest that the proposed approach has potential applications in search and rescue, environmental monitoring, and industrial automation.

 ISSN 1560-8956

Previous studies have identified ways to train individual objects as well as entire systems, and described the possibilities of data transfer between the modules of these systems, so it is advisable to use these methods and algorithms to study the possibility of training a system to search for objects in an unknown space and compare this algorithm with the usual search by individual objects without information exchange.

Problem statement

Suppose there is an environment with area S. For each experiment, we choose the number of objects n such that $S/n = const$, a value that is the same for all experiments. Let each experiment last for 200 time units, and for each time unit, the object can move at a constant speed equal to 1 unit of distance per time unit.

For reinforcement learning, let’s define the following parameters:

- for each unit of time until all objects are found, the system receives a penalty of 1 point;
- for each object found, the system receives a reward of 25 points.

If the system has found all the objects, the experiment ends.

For the experiment with combined objects, the search system receives the distances to each object as well as the coordinates of all other objects in the system, which will increase the computational time and potentially reduce the learning speed.

For a single-object experiment, the system receives as input the distance to unknown objects at each time point.

Main research material

We plan to conduct the following studies:

- in the first study, we will determine the effectiveness of the reinforcement learning algorithm for a given task with a system of search modules;
- in the second experiment, we will compare the two systems in an environment where the objects searched for by the system are stationary;
- in the third experiment, we will compare the same two systems, but this time the objects will move chaotically.

The size of the input data of the neural network is $n \cdot 2 + o \cdot 2$,

where n is the number of modules of the system,

o is the number of objects to be searched.

The size of the output data is n, the number of system modules, and each output element can take on 4 values – the direction of the next module step (up, down, left, right).

Having created the system according to the specified parameters, 15000 experiments were conducted for the system of combined modules. The result of the experiments is shown in Fig. 2, in the form of a dependence of the points obtained by the system on the number of experiments. From Fig. 2 shows that reinforcement learning is effective and the system works and improves its performance. It can also be seen that the system reaches its maximum value of 10 points per experiment. This is due to the need to overcome a certain number of steps to the object.

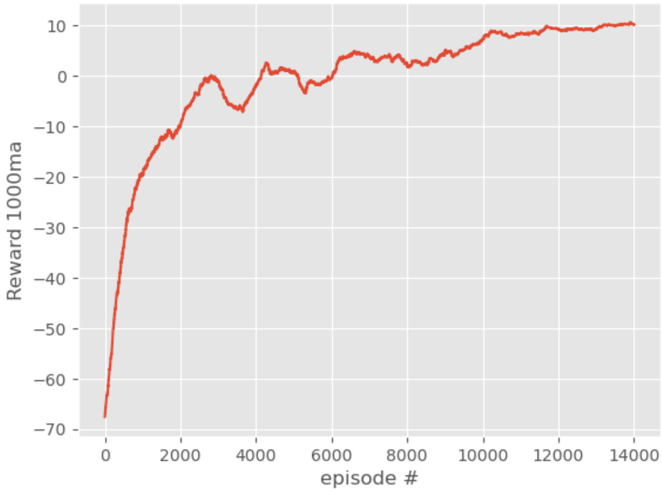

Figure 2. The number of points obtained by the system from search modules during 15000 experiments using reinforcement learning algorithms

The next experiment was a comparison of two systems, one from a separate object in a multiple smaller environment and the other with a system of combined modules. In Fig. 3, the first system is labeled System 1, and the second, respectively, System 2. The graphs show that the system of combined modules learns more slowly, but achieves a better maximum result.

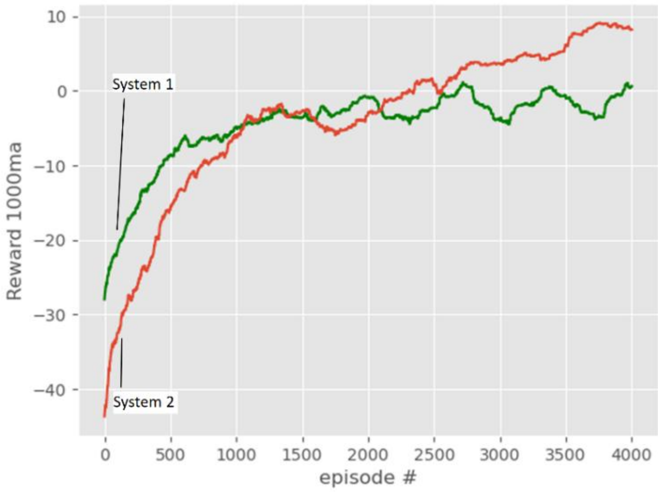

Figure 3. Comparison of the obtained scores for the experiment with a single object (System 1) and the system of combined objects (System 2) in the search for a static object

The next experiment was a comparison of the above systems for searching for moving objects. As can be seen from the graphs in Fig. 4, the system of the combined modules (System 2) in this experiment obtained the best performance both in terms of learning speed and the maximum score.

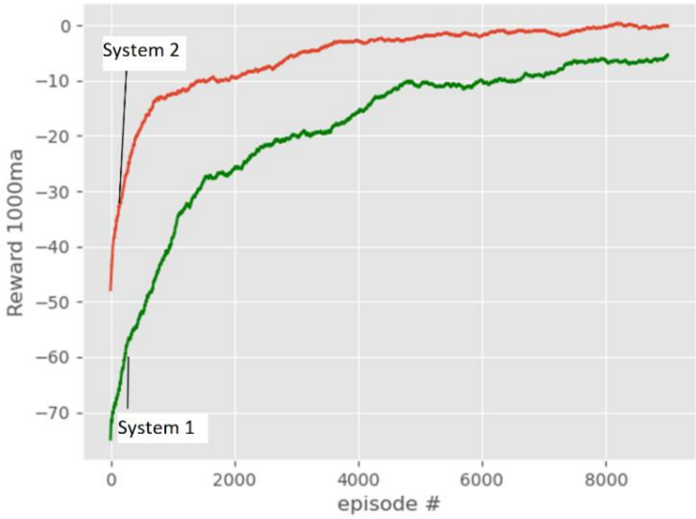

Figure 4. Comparison of the scores obtained for the experiment with a single object (System 1) and the system of combined objects (System 2) in the search for a moving object

The graphs show that the module system works more efficiently than using the same number of individual modules.

Conclusion

From the results obtained, we can conclude that using reinforcement learning, it is possible to obtain an algorithm for efficient search using a system of combined search modules that will be better than using the same number of search objects with separate control.

This research can be extended in several ways:

- by increasing the number of elements and the size of the environment, using more powerful computing power;

- by using a three-dimensional environment for experiments;

- using only the investigated area, environment for input data, making the environment itself unknown;

- adding other indicators (the amount of charge of the search object, the need for recharging, the possibility of malfunction, loss of the object, etc.)

REFERENCES

1. Playing Atari with Deep Reinforcement Learning/ Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, Martin Riedmiller // NIPS Deep Learning Workshop. – 2013.

2. Swarm modelling with dynamic mode decomposition / E. Hansen, S. L. Brunton, and Z. Song // Neural and Evolutionary Computing. – 2022.

3. A Scalable Reinforcement Learning Approach for Attack Allocation in Swarm to Swarm Engagement Problems / Umut Demir, Nazim Kemal Ure // Robotics. – 2022.

4. Distributed Machine Learning for UAV Swarms: Computing, Sensing, and Semantics / Yahao Ding, Zhaohui Yang, Quoc-Viet Pham, Zhaoyang Zhang, Mohammad Shikh-Bahaei // Machine Learning. – 2023.

5. Deep Reinforcement Learning for Swarm Robots / Maximilian Hüttenrauch, Adrian Sosić, Gerhard Neumann // Journal of Machine Learning Research. – 2019. – T.20.

 ISSN 1560-8956