PS-20 R&D TRENDS AND FUTURE TASKS IN THE PROBLEM OF CONTROL SYSTEM DEVELOPMENT FOR MARINE ENGINES USING A REINFORCEMENT LEARNING TECHNIQUE

環境・動力系 ＊ボンダレンコ オレクシー、福田 哲吾
流体性能評価系 北川 泰士
知識・データシステム系 澤田 涼平

1. Introduction

In recent years, the prevalence of machine learning techniques has spurred interest in the development of autonomous ships. Many studies have focused on the problem of collision avoidance, path planning, and steering control by applying a vast range of neural network techniques that can learn from humans' operation experiences – through supervised learning. Moreover, the recent progress in machine learning and computing power, especially deep learning, has opened the ways to develop systems that can learn from machines' own experiences – so-called reinforcement learning (RL). In general, there are three key parts constituting the RL algorithm: an environment which represents the behaviour of a system under consideration; a reward function which is directly related to the desired objectives and defines optimal and undesirable actions; an autonomous agent, encouraged by the reward function, attempts to learn the optimal action exploring the space of possible solutions until an appropriate policy (discrete set or continuous surface of control actions) that achieves the set objectives in the best possible way is found.

This report introduces the preliminary results of research and development carried out at the National Maritime Research Institute. In this research, the RL approach was applied to the control problems of marine engines. The first is the problem of the propulsion engine's optimal speed control considering the encountered sea conditions. The second is the problem of a marine gas-fueled engine's control considering a rapid transient response.

2. Optimal Control of Propulsion Engine

The speed control of the propulsion engine in the actual sea conditions is the key component for the safe and energy-efficient operation of the ship. In a rough sea, the efforts of the control system in stabilising shaft speed by continuously adjusting the engine torque, become undesirable ample. The primary consequence of such an operation is an unfavourable engine's thermal and dynamic state, resulting in increased fuel consumption. The remedy to such a situation is to reduce the efforts in adjusting the engine torque on account of increased engine speed fluctuation. Thus the trade-off control of the shaft speed can be achieved by readjusting the control system gains as shown in [1]. Notably, low gain and high gain are required for calm sea and rough sea conditions, correspondingly. In that respect, the task of RL is to train the control agent to properly select control system gains depending on the sea state with the objective of fuel efficiency. The overview of mutual relations in the RL framework is outlined in Fig. 1.

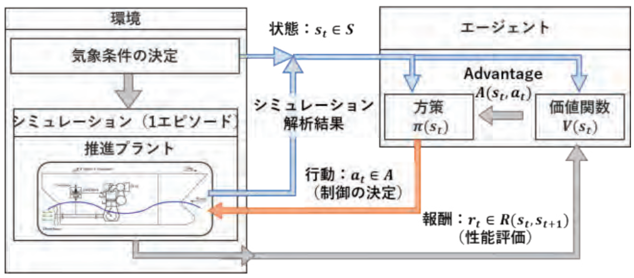

Fig. 1 Overview of the propulsion system RL framework

In the RL algorithm, an agent receives information from the environment and takes action. Thus, the realistic representation of the environment is indispensable, and for the purpose of this study, it consists of an engine torque generation model by fuel combustion, propeller torque model, and hull motion model. All elements are connected by the shaft rotation motion equation to form a propulsion system. On top of that is the engine speed governor model whose parameters are modified by the intended control agent. Details of a concrete mathematical model can be found in [2].

After performing the action, the agent receives a reward for its action and a new state of the environment. For the outlined task, the reward function should correspond to the improvement of fuel energy transformation efficiency from the combustion in the engine to the cargo transportation by a ship, considering also possible ship speed loss, δv_s.

Based on this consideration, the following reward function, R^t, is applied as a baseline design:

$$R^t = \sum_{i=1}^{n} r_i \, ; \, r_i = f(\dots, \delta v_s) \qquad (1)$$

The exact form of the constituent components is now under active development and thus can not be disclosed.

There are many frameworks for implementing RL algorithms. For the purpose of this study, the Mathworks MatLab RL toolbox was used, ensuring that any RL algorithms combination can be implemented quickly and conveniently in the developed environment. Figure 2 shows the example of RL progress. Although for robust learning the number of trial steps (episodes) should be of order 10^6, there is a tiny improvement in reward visible after 1000 trials.

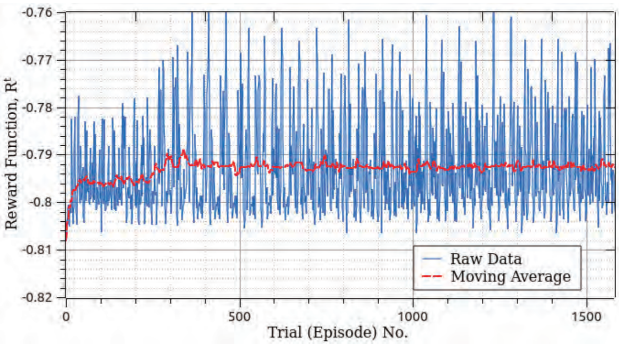

Fig. 2 Evolution of reward during the learning progress

3. Optimal Control of Gas-Fueled Marine Engine

Natural gas-fuelled engines are entirely worthy successors to diesel engines for maritime marine, especially regarding GHG emission reduction. However, the marine application implies transient operating modes, which require fast response and load fluctuation rejection. In that respect, the load acceptance of the gas-fuelled engines is subject to a specific limit due to a knock and misfiring phenomena. Thus for the efficient and safe operation, it is necessary to consider the control problem related to transient behaviour and develop countermeasures. In this context, the RL algorithm can be beneficial in solving the complex task of gas-fuelled engine control.

The environment consists of the previously developed gas-fuelled engine model as in [3], with two control loops: the first is controlling the amount of fuel to keep the engine speed, n_e, constant and the second is controlling the charge air pressure to keep the air-to-fuel ratio (*AFR*) constant. To this end, the task of RL is to train an agent which will assist the control loops in transient conditions and be rewarded for the simultaneous reduction of n_e and *AFR* deviations from the reference, denoted with ‘0’ subscript. Therefore, a reward function composed of mean square errors (MSE) is applied as a baseline design:

$$R^t = -\left[\sum_{i=1}^{N} \left(n_{e_0} - n_e\right)^2 + (AFR_0 - AFR)^2\right] \qquad (2)$$

Figure 3 outlines the implemented RL framework for the gas-fuelled engine. Figure 4 illustrates the example of RL training. As can be seen, the control agent provides simultaneous reduction of the reward function components below the standard value of MSE.

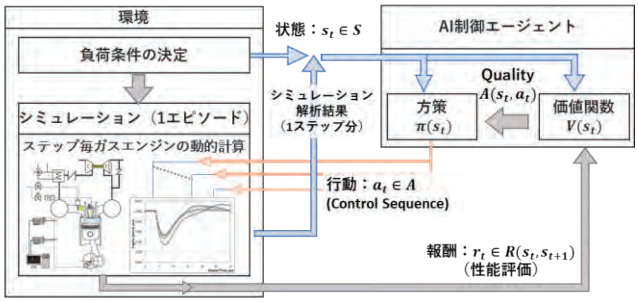

Fig. 3 Overview of the gas-fuelled engine RL framework

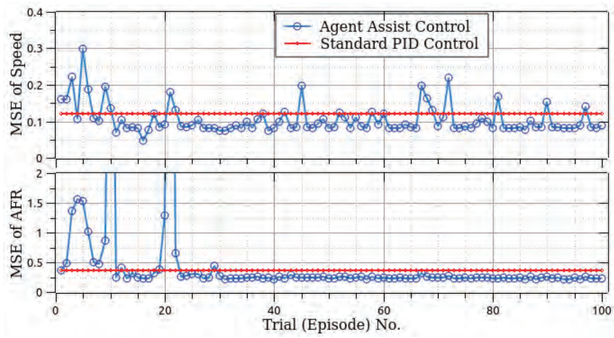

Fig. 4 Evolution of reward components in the learning

4. Conclusion

This report is just the beginning of the journey towards the application of the RL technique to the control problem of marine engines. At this stage, the realistic simulation environments for the RL algorithm have been created, based on past research activity. At the same time, there is still a question on the reward function design that affects the efficiency and the results of the RL to a great extent. We will continue to carry out basic research and contribute to the development of novel control techniques for marine engines.

5. References

1) Bondarenko O., Kashiwagi M.: Dynamic Behavior of Ship Propulsion Plant in Actual Seas, J. of JIME, Vol. 45, pp.76-81, 2010

2) Kitagawa Y., Bondarenko O., et.al.: A Fundamental Study on Fuel Consumption Performance in Waves of a Ship Installing a CFRP Propeller. J. of JASNAOE, Vol.35, pp.269-273, 2022

3) Bondarenko O., et.al.: Gas Engine for Marine Application: Modelling and Control. *28th CIMAC Congress*, 2016