CHAPTER 1

INDTRODUCTION

1.1 Background

Every year, more than 2 (two) million tons of plastic are dumped into rivers and eventually drift into the sea [1], so that the waste disposal system becomes a crucial sector [1][2]. Manual waste management methods are often used to overcome the crisis[3]. However, there are several problems with manual waste management, such as the safety of workers, not being able to reach remote areas, high operating costs, and others[3]. The Autonomous Trash Collector Robot (ATCR) is a solution to overcome these problems, because it can reduce the risk of accidents, can reach remote areas, and can do repetitive work [3, 4]. Autonomous mobile Robot has been developed in several studies, such as: Robot wall cleaning [5], water cleaner[6], and floor cleaner [4, 7, 8].

Automatic navigation on the robot is needed for environmental mapping so that the robot can run properly. Data from robot sensors can be mapped and used by robots for navigation and movement planning. In addition, data from sensors is also used to estimate the position of the robot needed when mapping the surrounding environment. Behavior-based robotic architecture is a control system that is not model-based, because it has a behavior structure that works together in parallel, concurrently and asynchronously (Brooks, 1986). In this approach, the system is broken down into several modules, each of which is responsible for performing a behavior (behavior). Each behavior contains a complete path from sensing to action. All modules that represent one behavior work together[2]. The more tasks the system, the more complex it is, so that it can lead to conflicts between behaviors. Therefore, a method of coordination between behaviors was developed. The main attention is paid to two approaches to coordination mechanisms, namely competitive/arbiter and cooperative/command fusion. In the competitive Coordinator method of ensuring controller robustness, only one behavior is allowed to provide a control signal. Meanwhile, cooperative coordination combines all existing behavior outputs and determines the performance of the robot's trajectory. The two main layers in the schema are the deliberative layers that divide the robot's mission into a set of tasks, and the behavior-based layer that is responsible for accomplishing these tasks. The thesis focuses only on the behavior-based layer. A behavioral coordination approach is proposed. Its main feature is the hybrid coordination of behavior, between competitive and cooperative approaches. In addition to the right architecture, it is also necessary to have the right learning mechanism on the robot to deal with the unexpected.

Reinforcement learning allows agents to select optimal behavioral policies through trial-and-error interaction training with the environment [1]. Recently, reinforcement learning has had great success in many tasks such as video games and control simulations [2]. Also, some robotics problems might naturally be expressed as one of reinforcement learning. Unlike other machine learning methods such as supervised and unsupervised learning, reinforcement learning provides feedback in terms of the behavioral environment of the interaction function to measure the one-step performance of the robot. In the process of learning reinforcement, the agent will finally conclude the optimal behavior policy after exploring the environment. Reinforcement learning (RL) can learn appropriate actions from environmental conditions. In the process of interaction between the agent and the external environment, the agent repeatedly learns through trial and error, relates environmental information, and continuously optimizes the agent's action strategy [3]. This optimization method gives RL an excellent decision-making ability [4]. Currently, learning reinforcement has been successfully applied in the path of planning mobile robots [3]. Reinforcement learning algorithms solve sequential decision problems posed as Markov decision processes (MDPs), studying policy by letting agents explore the effects of different actions in different situations while trying to maximize reward signals. RL has been successfully applied to various scenarios. Learning by demonstration is an approach to robots/learning agents taking demonstration inputs to construct an action or task.

Reinforcement learning is the optimal control method, when the agent starts from an ineffective solution which gradually increases according to the knowledge gained to solve successively. decision problems [4]. To use reinforcement study, several approaches are possible. The first consists of manually discrete issues to obtain state and action space; which can be used directly by the algorithm using table Q [4]. However, it is necessary to pay attention to discretization options, so as to allow true learning by providing situations and actions that contain understandable rewards. The second method consists of working on a continuous state and action space using a value function [5]. Indeed, to use reinforcement learning, it is necessary to correctly estimate the value function. The results obtained show a substantial improvement of the robot's behavior and learning speed.

In this study, hierarchical behavior architecture separates the primitive behavior tasks that coordinate with learning behavior [9]. Inspired by [4] research and research[10–13], this thesis research aims to design *Autonomous Trash Collector Robot* (ATCR) using *Reinfocement Learning* [**?**]. In this research, a robotic ROS middleware platform will be designed with a behavior-based control architecture. Robot (ROS) and Gazebo operating systems were used to simulate the virtual environment. Then Q learning will also be added as a robot learning mechanism. The robot will perform autonomous navigation to avoid obstacles and find targets.

1.2 Problem Identification

There are several problems with manual waste management, such as the safety of workers, not being able to reach remote areas, high operating costs, and others. Sensors are used for navigation and movement planning. In addition, data from sensors is also used to estimate the position of the robot needed when mapping the surrounding environment. The more tasks the system has, the more complex it is, so that it can lead to conflicts between behaviors

In addition to the right architecture, it is also necessary to have the right learning mechanism on the robot to deal with the unexpected. It is necessary to pay attention to discretization options, so as to allow true learning by providing situations and actions that contain understandable rewards. It is necessary to correctly estimate the value function

1.3 Objective

Implementing the Q-Learning algorithm and integrating it with the ROS Navigation Stack to avoid and find targets

1.4 Research Method

The Q-Learning algorithm on the mobile robot finds the optimal path without colliding with obstacles. To complete the Q-Learning model that takes sensory data given to the robot as input (or observation), the output of the model is instructions for the robot's movements, such as moving forward, turning left, turning right, and back, which are executed by modifying the robot's wheel position. mobile virtual robots and environments for navigation tasks using Gazebo [11], which provides a multi-robot 3D environment in an open-source format.

1.5 Scope of Work

1. The location where the test is performed and uses an artificial environment.
2. There are randomly stored objects and obstacles to avoid.
3. Robot only goes to the specified destination point. Some advanced functions such as fetching objects, transporting objects, and other functions are ignored.
4. The algorithm training process is carried out through simulation.

1.6 Metodologi

The stages in the research are as follows:

1. **Study Literature** navigation mobile robot, Behavior based Robotic, Reinforcement Learning, Q-learning, Robot Car Simulation method from previous research. such as in journals, written works, as well as books and reference literature related to research.

2. **Requirements Analysis** At this stage an analysis is carried out which includes the need to conduct research, The analyzed requirements are divided into data analysis and system requirements analysis. The analysis is carried out so that the system built can run according to the previously determined design.

3. **System Design** Doing software design after studying literature study and state of the art mobile robot navigation system. and make the structure of the reinforcement learning algorithm, based on the research topic problems, and perform simulations on the system designed.

4. **Evaluation** Implement and test the system, training the reinforcement learning algorithm used, and test the software that has been designed, then evaluate the system based on research topics.

5. **Final Test, Result and Analysis** the process of collecting data and results in accordance with research objectives, system analysis of the results of system testing.

6. **Reporting** Writing the final report of research and journal publications.

1.7 Systematic Writing

The systematics of writing this thesis are as follows:

1. CHAPTER 1 In this section is the background of the problem, problem formulation, objectives, research method, methodology, and systematics of writing.

2. CHAPTER 2 In this section, we review the literature review related to the theory, technology and methods used in mobile robot navigation systems using ROS and reinforcement learning.

3. CHAPTER 3 This section discusses the design of a mobile robot system on the ROS platform with reinforcement learning into navigation design.

4. CHAPTER 4 This section discusses testing an integrated navigation system using reinforcement learning and describes the test results.

5. CHAPTER 5 This section includes conclusions from the results and analysis of the thesis research as well as suggestions for the development of further research.