

|  |
| --- |
| School of Computing  Faculty of Engineering |

Control System (AI) for Wrestling Robot

Fanhui Meng

Submitted in accordance with the requirements for the degree of  
MSc Advanced Computer Science (AI)

**Session 2019/2020**

The candidate confirms that the following have been submitted*:*

*<As an example>*

|  |  |  |
| --- | --- | --- |
| **Items** | **Format** | **Recipient(s) and Date** |
| *Deliverables 1* | *Report* | *SSO (xx/xx/xx)* |
| *Deliverables 2* | *Code and URL* | *SSO (xx/xx/xx)* |
| *Deliverable 3* | *Youtube video URL* | *Supervisor, assessor (xx/xx/xx)* |
| *Deliverable 5* | *User manuals* | *Client, supervisor (xx/xx/xx)* |

Type of Project: Exploratory Software

The candidate confirms that the work submitted is their own and the appropriate credit has been given where reference has been made to the work of others.

I understand that failure to attribute material which is obtained from another source may be considered as plagiarism.

(Signature of student)

© 2020 The University of Leeds and Fanhui Meng

# Summary

The artificial intelligence technology has developed for many decades and used in many fields. This project is focusing on using AI technology, design a high-level control system for Zumo robot in the Sumo robot league.

The Sumo robot league is a popular robot competition, and the main rule of this competition is two vehicle-like robots without mechanical arm trying to push each other out of the ring. The Zumo robot is the one that is going to use in this project. It's an off the shelf Arduino-based robot. Thus, no structure design or hardware design (e.g. circuit design) or low-level design (e.g. PWM motor speed control) in this project.

The high-level control system means the ‘brain’ of the robot. It has strategies to cope with different situations. There are many machine learning methods to develop strategies for the robot. Basically, there are two kinds of machine learning can be applied in this project, which is the supervised learning and unsupervised learning. One is to develop the control system by telling the robot, which is the right thing to do and what is wrong. Another is to develop the control system by not telling what the robot should do, but only reward or punishment the robot according to the rules and it’s behaviours.

This project aims to implement and compare two different control systems, one is supervised, and another is unsupervised.

Furthermore, this project will do more research about how the same control systems or the same ideas can be applied in other robot competitions or the robot in daily life (e.g. sweeping robot).

# Acknowledgements

aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

*<This page should contain any acknowledgements to those who have assisted with your work. Where you have worked as part of a team, you should, where appropriate, reference to any contribution made by others to the project.*

*Note that it is not acceptable to solicit assistance on ‘proof reading’ which is defined as “the systematic checking and identification of errors in spelling, punctuation, grammar and sentence construction, formatting and layout in the text”; see* [*http://www.leeds.ac.uk/qat/documents/policy/Proof-reading-policy.pdf*](http://www.leeds.ac.uk/qat/documents/policy/Proof-reading-policy.pdf)*. >*

# Table of Contents

Summary iii

Acknowledgements iv

Table of Contents v

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Overview 1

1.2 Aim and Objectives

Chapter 2 Tables and Figures 2

2.1 Tables using the ‘table caption’ and ‘table description’ Styles 2

2.2 Figures using the ‘figure caption’ and ‘figure description’ Styles 2

List of References 3

Appendix A External Materials 4

A.1 Level 2 Heading with ‘heading 2’ Style Applied by Pressing Ctrl Shift 2 4

A.1.1 Level 3 Heading with ‘heading 3’ Style Applied by Pressing Ctrl Shift 3 4

A.1.1.1 Level 4 Heading with ‘heading 4’ Style Applied by Pressing Ctrl Shift 4 4

Appendix B Ethical Issues Addressed 5

B.1 Level 2 Heading 5

# List of Figures

Figure 1.1 Main features of the Zumo 32U4 robot

Figure 2.1 Decision tree example for surviver on Titanic (Wikipedia Contributors, 2019)

Figure 2.2 General process for GA

Figure 2.3 General procedure for reinforcement learning

# List of Tables

Table 2.1: Fuzzy rules for sumo robot (Erlan et al., 2013)

Table 2.2: Genes for robot attributes (Lehner et al., 2019)

Table 2.3: Bayesian example for sumo robot

# List of Abbreviations

BEAST Bio-inspired Evolutionary Agent Simulation Toolkit

GA Genetic algorithm

FLC Fuzzy Logic Control

# Chapter 1 Introduction

## Project Aim

The aim of this project is to design a fine robot high-level control system, which is the ‘brain’ of the non-arm Zumo robot. The control system would have its own strategy and drive the robot's movement during the competition. As for the strategy, it can be assigned with specific movement in different cases, or the robot can develop its own strategies, which may be related to the evolutionary algorithm. And this will be discussed later in the report.

The main goal of this project is to make the Zumo robot be competitive and perform well in the Sumo robot competition. Besides, this project will find out if the idea of a control system can be applied to a wider range of different robots, such as other robot competition or the robot service in daily life.

This project is going to use the Zumo 32U4 robot, which is a complete, versatile robot controlled by an Arduino-compatible Atmega32U4 microcontroller. Therefore, extra hardware structure and improvement is not considered in the project. The Zumo robot has two motors, one Atmega32U4 chip as the brain and a variety of sensors, including two front proximity sensors, two side proximity sensors, five line sensors (three of them are the receiver, two of them are emission) and accelerometer. So the Zumo robot can detect the opponent and run towards or away from it, which satisfy every requirements of a robot in the Sumo league.



Figure 1.1: Main features of the Zumo 32U4 robot

Sumo robot league is a very popular international robot wrestling competition, which is two robots attempt to push each other out of the ring. Thus, the wrestling robot must be capable of autonomously locating the opposing robot and pushing it out of the ring (Wilson et al., 2016). The last stand robot in the ring is the winner. Also, there are other rules in the Sumo league, which the control system should be designed according to these rules. And these rules will be explained in details in the next chapter.

### Objectives:

* To get familiar with the hardware functions of Zumo robot.
* Conduct a theory study to compare different control system in the wrestling case.
* Implement two or more different control system and compare it’s advantages and disadvantages.
* Create a wrestling simulation environment. (Due to the lockdown policy, it’s hard to find the opponent in the real world)
* Evaluate the results with its performance in different simulation cases.

#### Deliverables

1. A software that can simulate the Sumo robot wrestling. Built using BEAST.
2. A Github repository that contains the source code of the system.
3. A developer documentation that provides:

* An overview of the simulation, algorithm used, programming languages and style.
* Instructions for setting up the project in a local development environment. (Provide VM for Mac or Windows user)

1. The MSc project report.

#### Ethical, legal, and social issues

Need to be done.

No Ethical, legal and social issues related to this project.

# Chapter 2 Background Research

## 2.1 Literature Survey

The sumo robot contest feature two self-controlled robots trying to push each other out of a ring. (Anon, n.d.) The basic rules are the first robot that touches outsides of the ring loses the round, and the last stand robot wins the round. The match is the best of three sets. And the robot wins the most matches are winning the contest.

Several structure requirements for the participant robot (e.g. robot size, mass, etc.) are clearly defined in the sumo robot rules. But these requirements are not being considered in this report, since this project's experiment and simulation are based on ready-made Zumo robot, which already satisfies all the hardware and structure aspect requirements.

About the contest environment. It’s a large, flat ring. It’s made of smooth, rigid wood. The top surface is dull black, and all of these black areas are in bounds.

A robot is usually started by pressing a button or other ways, such as hand-clapping, a whistle and so forth. After start command, the contestants immediately leave the exterior area around the ring. During the round, all people and object must be kept out of the ring and exterior area to avoid distracting the robots or altering the outcome. Upon pressing the start buttons, each robot must not move at all for five seconds.

A robot loses a round when any portion (including touch sensors, whiskers, scoops, or skirts) of the robot touches outside the ring. It doesn’t matter if the robot falls out on its own or is puhsed out.

If the match satisfiy any of the conditions below, the reference may choose to restart a round.

* Three minutes have expired
* No progress has been made in some period of time
* The robots fail to touch each other for some period of time
* The robots are hopelessly entangled or otherwise deadlocked
* Both robots fail to start or both contestants signal stoppage

This project design is based on the sumo rules above.

There are pretty much studies in Sumo robot control, and even more research on the general intelligence control system. And there are also different aspects of these project. Some are fucus on the low-level development, some are meant to design the hardware or the construction of the Sumo robot, some are focus on the electronics and so on. And this project is going to focus on the high-level design.

One study is to develop the Neuro-Fuzzy (NF) hybrid system as the control system (Erdem, 2011). In other words, it uses two systems, which is ANN and Fuzzy Inference System (FIS). FIS is for detecting and tracking the opponent, which relates sensor output (IR sensor) to motor control pulses. ANN is used for rule extraction and tuning the FIS parameters. And the result shows that this control system can improve the robot responses during competition. This is a good thought to develop a good control system. It's pretty much the low-level development. It's just like human work out to improve physical strength, but I would focus on brain development. However, the wrestling environment is uncertain, and the data is non-linear. Thus, the method that using ANN to improve fuzzy control is a good thought for high-level control as well.

Similarly, one project also uses fuzzy logic as the main idea of sumo robot control. It uses single fuzzy logic control (FLC) as the microcontroller for detection and tracking of an opponent in the competition ring. Three infrared sharp sensors for target detection. Then the fuzzy controller fuses the sensor data and provides the control signal to the motor for driving the robot toward the opponent (Erdem, 2007).

Then the author develops the fuzzy rules, which are intuitive rules that can be driven by all possible scenarios with input sensor values. For example, one of the fuzzy rules is that, if S1 and S2 detects the weak signal and S3 detects medium signal, the logical action of the robot is to control the motor to take a small turn to right. The optimized rules for detection and tracking of target are shown in below.



Table 2.1: Fuzzy rules for sumo robot (Erdem, 2007)

For every sensor, there are three fuzzy sets, including Low, Medium, High. The five singleton membership functions are Full Left, Small right, etc. Based on the output of the fuzzy controller, the motor will make the robot turn left or right.

Besides, a few studies use genetic algorithm (GA) to optimize the sumo robot. One study uses Java with Eclipse and the *dyn4j physics engine* for simulating the sumo robot fight (Lehner et al., 2019). Each robot has six attributes (e.g. the speed, search range), which are allocated to gene positions and can be controlled by the GA. The intensity of each genetically controlled capability is the characteristics, which is expressed with a value ranging from 1 to 9.

The genes for robot attributes and it’s controlled capability is shown below.



Table 2.2: Genes for robot attributes (Lehner et al., 2019)

The process for each individual robot is, at the start of the fight, each robot receives a randomly selected mix of values, which is the mix characteristics for the robot. Then run the fight to determine the highest fitness. The fitness is defined as how many movements are required until it's lose or stands last. The next step is the crossover and selection for the next generation. This project uses *roulette wheel selection*, which distributes the probability for the selection of each robot based on their relative probability of the fitness. And 50% mutation rate per gene is also applied to increase the probability of changing the characteristics randomly. Finally, create a new generation and transfers the new generation in a new loop. After thousands of generations, the last robot stands in the ring is the final winning genes of the GA optimization.

## 2.2 Methods and Techniques

2.2.1 Fuzzy Logical Control System

Fuzzy control is an intelligent control method based on fuzzy set theory and fuzzy logic inference. It's an intelligent control method that imitates a human's fuzzy inference and decision-making process from the behaviour. The method first compiles the experience of the operator or expert into fuzzy rules, then fuzzifies the real-time signal from the sensor. Use the fuzzed signal as the input of the fuzzy rule, completes the fuzzy inference and output the inference.

Several things are important for defining a fuzzy controller. First is the fuzzy interface, which is for performing a conversion from a certain real value of the input into a fuzzy set. For example, in this project, it may transfer the proximity sensor output, such as a number value, to a fuzzy set, like low, medium or high. Secondly, the knowledge base, which is consist of data base and rule base. The data base stores all the input and fuzzy set. While the rule base is the fuzzy rules usually based on expert knowledge or personal experiment. Fuzzy rules are normally connected by a series of relational words, and the most commonly used relational words are if-then, else, etc. Finally, is the inference and defuzzy-interface. The result of the inference indicates that the rule inference function of fuzzy control has been completed. However, the result obtained still cannot be used directly as a control variable. A conversion must be made to have a clear control variable output, which is the defuzzification. In this project, this could be used for determination of real value for directing of motors or other controls.

Fuzzy logic control doesn't require an accurate mathematical model. However, the factors such as the integrity of fuzzy control rules, the definition of fuzzy subsets and the fuzzy inference mechanism will have an impact on the performance of the fuzzy controller. Thus, most of the factors depend on the experience of experts. For this project, the main basis of control is based on a large amount of personal experience, knowledge and strategy in the sumo league.

2.2.2 Decision tree based Control System

Similarly to FLC, the performance of decision tree based control system also depends more on personal experience and strategy in a wrestling robot. As one kind of supervised learning, decision tree can be well applied in the classification as well as regression problem.

The decision tree algorithm uses a tree structure and uses layers of reasoning to achieve the final classification. It’s consist of one root node, leaf node and internal node. Root node contains full set of samples, internal node corresponding characteristic attribute and the internal node represented the final result or decision.

For example, in the figure 2.1 below. The root node is the ‘gender’ node, which must be the purest feature. ‘age’, ‘sibsp’ are the internal node. ‘survived’, ‘died’ are the predicted result, which is the leaf node.



Figure 2.1: Decision tree example for the survivor on Titanic (Wikipedia Contributors, 2019)

Basically, there are three-step to build the decision tree. First is the feature selection, and it's determins which features are used to make judgements. In the training data set, there may be many attributes of each sample, and the effects of different attributes are different. Therefore, the function of feature selection is to screen out the features that are more relevant to the classification results. In short, it's intended to find out the features with strong classification ability. After the features are selected, the next step is to generate the decision tree. It's triggered from the root node, calculated purity for the node. The feature with the most purity is selected as the node feature, and the child nodes are established according to the different values of the feature. Each child node is generated in the same way until purity is the lowest or there are no features to choose from.

The key things here is the word ‘purity’. It’s decide which is the optimal partition feature. And it’s means if most of the sample that contained in one tree node belongs to one category, then we say this tree node is pretty pure. There are mainly two ways of representing the purity, which is the information entropy and gini index. According to this, three decision tree algorithm is introduced, including ID3, C4.5 and CRAT. ID3 is based on information gain to choose feature. C4.5 is quite similar to ID3, which use information gain ratio. And CRAT use gini index to choose feature, which can be used in both classificaiton and regression problem.

After generating the decision tree, the final step is an option, which is pruning. The main purpose of pruning is to prevent overfitting by removing unnecessary branches.

Furthermore, random forest is another possible solution for the control system, which is made of many independent decision trees. When performing a classification task, new input samples are entered, and each decision tree in the forest is judged and classified separately. Each decision tree will get its own classification result. And the result that most of the decision tree classified is the final result for the random forest.

2.2.3 Artificial neural network

An artificial neural network (ANN) is a computing model, which is composed of a large number of artificial neurons connected to each other. Each node represents a specific output function. Each connection between two nodes represents a weighted value for a signal passing through the connection, which is called the weight. The network itself is usually an approximation of a certain algorithm or function or an expression of a logic strategy.

Feedforward neural network (FNN) is one kind of ANN. It's the simplest and widely used type of ANN. In this network, the information moves in only by forward. There are no cycles or loops in the network.

For this project, the sensor value can be the input of the network and output is the generated strategy. Moreover, the ANN’s parameter (e.g. weight, basis) can be adjusted or optimized or evolved through training. Thus, different way of optimization can be applied, such as Neuro-fuzzy or GA with FNN.

2.2.4 Naïve Bayesian Model

Bayesian is a kind of classificaiton method. For example, if a robot’s right proximity sensor detect obstacle while the left IR detect nothing, then this situation can be classify as obstacle in the right hand side of the robot. The content of the classification algorithm is to require a given feature to let us get the category, which is also the key to all classification problems. And how to specify the characteristics to get our final category is the key thing to be considered.

One example for sumo robot can be shown as below. For proximity sensor, 0 means the distant to obstacle is close, 1 means the distant is far, 2 means too far. For line sensor, 0 means out of ring, 1 means in the ring.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Proximity Sensor  Front Right (S1) | Proximity Sensor  Front Left (S2) | Line Sensor  Right (S3) | Line Sensor  Left (S4) | Label |
| 0 | 0 | 1 | 1 | Enemy Front |
| 0 | 1 | 1 | 1 | Enemy Front Right |
| 0 | 2 | 1 | 1 | Enemy Front Right |
| 1 | 0 | 1 | 1 | Enemy Front Left |
| 1 | 1 | 1 | 1 | Enemy Front |
| 1 | 2 | 1 | 1 | Enemy Front Left |
| … | … | … | … | … |

Table 2.3 Bayesian example for sumo robot

The bayesian formula is . A means the features, and B means the label or category. And each feature should be independent.

So in the sumo case example, if it require to calculate the possibility if the enemy is in the front or left or right, and the sensor values are S1, S2, S3, S4 respectively. Then use bayesian formula to calculate the possibility of front or left or right respectively.

First, calculate the possiblity of enemy in the front. It can be represented as below.

Because of each feature is independent, .

And

After getting everything we need, can be calculate.

Then do the same to and . The posibliity with the largest value is the final predicted result.

In general, the good thing about bayesian is that it’s easy to implement and supposed to have less error rate.

2.2.5 Genetic Algorithm

Genetic Algorithm (GA) is based on Darwin’s theory of evolution, simulating natural selection, natural selection, survival of the fittest. The main purpose of GA is to keep the best gene so that it will evlove better and more fit gene or individuals.

GA normally uses a fixed-length linear binary representation for it’s genotype. It’s can be designed as containing the control strategy of the sumo robot. And it’s a heuristic algorithm, which is suitable for Non-deterministic polynomial (NP) problem. This will be explained in the next section.

The general steps of GA is shown below.



Figure2.2: General process for GA

At the first generation, the population will be the group of randomly generated solution. Then use the fitness function to calculate the fitness of each individual separately.

The fitness is the most crucial thing of GA. It’s play the role of ‘God’ in GA, which measure the pros and cons of individual and decide who is going to stay or being eliminated. For example, in sumo cases, the fitness can be if the robot win the most matches, then it’ll get more fitness score, while the one lose the most matches are tend to be eliminated. GA will perform N generations and each generation will generate several individuals, which is population. The fitness function will give a score to all the individual to judge the fitness for each individual. Only the individual with higher fitness are retained, so the quality of population will become better and better after several iterations.

After calculating the fitness score, the next step will be the selection, which is also quite important for GA. There are several ways of selection. For example, the main idea of roulette wheel selection is that individuals with better fitness values are more likely to be selected. The tournament selection takes a certain number of individuals from the population each time, and selects the best one to enter the offspring population. Repeat this operation until the new population size reaches the original population size. Ranked based selection, only select the top individual according to fitness score. There are also many other selection method, like stochastic tournament, excepted value selection, truncation selection, and so on.

After selection, GA will finally going to evolution process. Do the crossover first, then mutation. The crossover need to find two chromosomes from the selected individuals of the previous generation. Then cut a certain position of the two chromosomes and splice them together to generate a new chromosome. This new chromosome contains a certain number of the two individuals’ genes from the last generations. Mutation simply means to change the small part of the genes randomly. Introduce mutation can help the solution escape from local optimal, and is helpful for the algorithm to find the global optimal solution.

Then calculate these new generations’ fitness score, and do the selection, crossover, mutation again until the result is good enough or reach the number of iterations limits.

2.2.6 Reinforcement learning

Reinforcement learning is also an unsupervised learning. And it’s usually described by Markov Decision Process (MDP). Then overall concept is for reinforcement learning is that, a macine is in an environment and it can only interact with environment through actions. And the state is the machine’s perception of the current environment. When the machine performs an action, the environment will feed back to the machine a reward based on the potential reward function. Also it will make the environment transfer to another state according to a certain probability. In summary, reinforcement learning mainly includes four elements, which are state, action, transition probability and reward function.

The general procedure for reinforcement learning as below.



Figure 2.3 General procedure for reinforcement learning

As agent performs a certain task, it first interacts with the environment to generate a new state, and the environment gives a reward at the same time. Then continue the iteration, the agent and the environment continue to interact to generate more new data, and modify its own action strategy accordign to these new data. After several iterations, the agent will learn the action strategy needed to complete the task.

In this project, the agent would be the Zumo robot. And state including the environment state, which is the sumo ring, agent state, which is the input data for agent and information state, which include useful information needed for future prediction. Environment is the partially observable environment, which means the agent will discover the environment by with it’s sensor.

The algorithm for reinforcement learning including Q learning, Sara, Deep Q network and so on. For example, the famous Q learning is based on Q chart like the chart below.

|  |  |  |
| --- | --- | --- |
| Q | A1 | A2 |
| S1 | -3 | 3 |
| S2 | -4 | 4 |

Table 2.4 Example for Q chart

S means state, A means action in Table 2.4. In this Table 2.4, , which means A2 rewards more than A1. Thus, it would be reasonable to choose A2 as the action. Then comes to S2, repeat the previous steps, choose the most rewards action and get to the next state.

However, the Q chart include the estimated value, so this chart needs to be updated. For example, , so multiply the max value by an attenuation value gamma (e.g. 0.8) and add the reward when reaching S2. And this value be seen as the actual value. Now we got the actual and estimated value, the Q chart is able to be updated. According to the difference between estimation and actual, multiply this difference by a learning rate and add the old estimation value , then got the new update value.

## 2.2 Choice of methods

Decision tree.

GA with FNN.

pass

# List of References

Anon n.d. Autonomous Sumo Robot Rules - Robot Room. *www.robotroom.com*. [Online]. Available from: <https://www.robotroom.com/SumoRules.html#:~:text=Mini%2Dclass%20Sumo%20robots%20may%20be%2010%20centimeters%20(3.93%20inches>

Erdem, H. 2007. A Practical Fuzzy Logic Controller for Sumo Robot Competition. *Springer-Verlag Berlin Heidelberg 2007*. (LNCS 4815), pp.217–225.

Erdem, H. 2011. Application of Neuro-Fuzzy Controller for Sumo Robot control. *Expert Systems with Applications*. **38**(8), pp.9752–9760.

‌ Erlan, C., Lima, O., Almeida De Araújo, F., Bibiano Da, M., Júnior, S., Edson, A., Filho, R., De Andrade, R., Rabêlo, L., Allisson, T., Da Silva, R., Jose, A. and Alves, O. 2013. *An Enhancement in Conventional Potential Field Using a Fuzzy System for Navigation of a Sumo Robot*.

Lehner, J., Dornberger, R., Simić, R. and Hanne, T. 2019. *Optimization of Multi-Robot Sumo Fight Simulation by a Genetic Algorithm to Identify Dominant Robot Capabilities*.

Wikipedia Contributors 2019. Decision tree learning. *Wikipedia*. [Online]. Available from: https://en.wikipedia.org/wiki/Decision\_tree\_learning.

Wilson, B., Author, S., Germann, T. and Al-Olimat, K. 2016. *Sumo Robot Competition* [Online]. [Accessed 4 August 2020]. Available from: http://people.cst.cmich.edu/yelam1k/asee/proceedings/2016/student\_regular\_papers/2016\_asee\_ncs\_paper\_58.pdf.

# Appendix A External Materials

<Level 1 Heading with ‘heading 1’ Style Applied by Pressing Ctrl Shift 1> Text under appendix heading. Text under appendix heading. Text under appendix heading. Text under appendix heading. Text under appendix heading. Text under appendix heading.

## A.1 Level 2 Heading with ‘heading 2’ Style Applied by Pressing Ctrl Shift 2

Text under level 2 heading. Text under level 2 heading. Text under level 2 heading. Text under level 2 heading.

### A.1.1 Level 3 Heading with ‘heading 3’ Style Applied by Pressing Ctrl Shift 3

Text under level 3 heading. Text under level 3 heading. Text under level 3 heading. Text under level 3 heading.

#### A.1.1.1 Level 4 Heading with ‘heading 4’ Style Applied by Pressing Ctrl Shift 4

Text under level 4 heading. Text under level 4 heading. Text under level 4 heading. Text under level 4 heading.

# Appendix B Ethical Issues Addressed

Text under appendix heading. Text under appendix heading. Text under appendix heading. Text under appendix heading. Text under appendix heading. Text under appendix heading.

## B.1 Level 2 Heading

Text under level 2 heading. Text under level 2 heading. Text under level 2 heading. Text under level 2 heading.