Design of Multiple-Functions Controller Based on Machine Learning

Graduate School of System and Information Engineering
University of Tsukuba

March 2015

Syafiq Fauzi Bin Kamarulzaman
# Contents

Abstract iii

1 Introduction 1
  1.1 Research Background 1
  1.1.1 Multi-Functionality 1
     1.1.1.1 Multi-Functionality against Non-Linearity 2
     1.1.1.2 Embedding Human Knowledge for Multi-Functionality 4
  1.1.2 Learning Control 5
     1.1.2.1 Reinforcement Learning 5
     1.1.2.2 Absence of Multi-Functionality in Reinforcement Learning 7
  1.2 Research Objective & Contents 9
     1.2.1 Research Objective 9
     1.2.2 Research Content 10

2 Multiple-Functions Learning Control by Substitute Target 12
  2.1 Substitute Target 12
  2.2 Utilization of Substitute Target 14
     2.2.1 Control Function for Substitute Target System 15
     2.2.2 Recognition Function for Substitute Target System 17
     2.2.3 Learning Function for Substitute Target System 19
  2.3 Experiments & Results 21
     2.3.1 Experiments Settings 21
     2.3.1.1 Inverted Pendulum Model 21
     2.3.1.2 Parameters for Control System Implementation 23
     2.3.1.3 Simulation on the Usage of Substitute Target Knowledge 25
2.3.1.4 Simulations on Learning Control through Direct Control Constraints ........................................ 26
2.3.1.5 Simulations on Learning Control through Direct and Indirect Control Constraints ................... 28
2.3.2 Experiments Results .................................................................................................................. 30
  2.3.2.1 Experiments Results by Simulations .................................................................................... 30
  2.3.2.2 Experiments Results by Real Machine Operation ................................................................. 36

2.4 Summary ..................................................................................................................................... 37

3 Multiple-Functions Learning Control by Multiple Control Knowledge ............................................ 38
  3.1 Multiple Control Knowledge in Learning Control ........................................................................ 38
  3.2 Application of Multiple Control Knowledge in Learning Control: Rapid Position Control ............ 40
    3.2.1 Introduction to Rapid Position Control .................................................................................. 40
    3.2.2 Manipulation of Angular Orientation for Rapid Position Control ........................................ 42
      3.2.2.1 Emulating Angular Orientation Control of an Aerial Hovering Vehicle on Pendulum System 42
      3.2.2.2 PD Control of Angular Orientation on a Cart Pendulum System ...................................... 44
      3.2.2.3 Obtaining Rapid Position Control Using Angular Orientation for Inverted Pendulum ........ 46
    3.2.3 Simulation Settings ................................................................................................................ 47
    3.2.4 Simulation Results ................................................................................................................. 49
      3.2.4.1 Knowledge Improvement through Learning Process ...................................................... 49
      3.2.4.2 Successful Operation Learned through Simulation ......................................................... 50
  3.3 Application of Multiple Control Knowledge in Learning Control: Rapid Position and Obstacle Control 52
    3.3.1 Parameters of Learning Control for Rapid Position and Obstacle Control .............................. 52
    3.3.2 System Structure for Rapid Position and Obstacle Control .................................................... 54
    3.3.3 Simulation Settings ............................................................................................................... 55
    3.3.4 Simulation Results ................................................................................................................. 57
List of Figures

1.1 States for control of Cart Pendulum System. ........................................... 2
1.2 Cascade PD control of a Cart Pendulum System. ................................. 3
1.3 Example of aerial hovering vehicle with non-linearity. (Parrot inc.) ....... 3
1.4 Control operation during position transition of aerial hovering vehicles. ... 4
1.5 Structure of functions in an Intelligent Control System. ..................... 4
1.6 Interaction between policy, reward function and value function. ........... 6

2.1 Substitute target provides continuity of action by providing an intermediate state. ................................................................. 13
2.2 Substitute target can be rearranged to satisfy the need for successful control manoeuvre along constraint states. .............................. 14
2.3 System structure for substitute target application. ............................. 15
2.4 Swing and stabilization control of the cart-pendulum system. ............... 16
2.5 The swing control of the pendulum based on substitute target. ............... 17
2.6 State clusters created from pendulum angle and pendulum angular velocity. 18
2.7 Constraints of the cart and the pendulum. ........................................... 19
2.8 State and action relation for substitute target based Q-learning. .......... 20
2.9 Control processes assigned according to pendulum angle. ...................... 20
2.10 Diagram of the cart-pendulum parameters. ........................................ 21
2.11 Cart-pendulum device (Japan E.M. Co., Ltd.) on which the simulations were based. ................................................................. 24
2.12 Movement range of the cart to satisfy the knowledge limit. .................. 25
2.13 The substitute target knowledge used as initial state. ........................ 26
2.14 Constraints assigned among the cart position for simulations. ............... 27
2.15 Constraints arranged around the cart position and the pendulum angle. ... 29
2.16 The average control success rate of the swing control for the first subject of the simulation. ................................... 30
2.17 The updated knowledge for both constructed knowledge and random knowledge after the simulation for Learning Control System by substitute targets. 31
2.18 The average success rate for pendulum swing control among direct control constraints. ................................................................. 32
2.19 Cart movement during the successful swing control for simulations for the three assigned cases of direct control constraints. .............................. 32
2.20 Comparison between initial substitute target knowledge and final substitute knowledge for each simulation concerning the direct control constraints. 33
2.21 The average success rate from every 10 trials for Pendulum Swing control among direct and indirect control constraints. .......................... 34
2.22 Comparison between initial substitute target knowledge and final substitute knowledge concerning direct and indirect control constraints. 35
2.23 Success area and constraint area detected during the simulations with direct and indirect constraints. ............................................................ 36
2.24 Cart movement during the successful swing control for simulation P4M4 and result from application on a real machine. ................................. 37

3.1 Structure of Learning Control System by multiple control knowledge. .... 39
3.2 Configuration of aerial hovering vehicle by angular orientation. .......... 41
3.3 Position control of an aerial hovering vehicle using target angle $\theta_T$ as reference. 41
3.4 The stabilization control of inverted pendulum. ................................. 42
3.5 The position control of inverted pendulum using target angle $\theta_T$ as reference. 43
3.6 Structure of system with multiple control knowledge for rapid position control. 43
3.7 PD control of cart-pendulum system emulating aerial hovering vehicle. . 44
3.8 The reference data used to calculate the preservation period of output for each target angle. ................................................................. 45
3.9 The structure of Learning Control System by multiple control knowledge for rapid position control of aerial hovering vehicles. ................. 46
3.10 Target position assigned for simulation of rapid position control. ........ 48
3.11 Improvement of the final cart position with respect to the number of trials. (Target position, $x_T=0.5[m]$) ........................................... 49
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.12</td>
<td>Angular trajectory of the pendulum during control operation that uses control knowledge obtained after 550 trials.</td>
<td>50</td>
</tr>
<tr>
<td>3.13</td>
<td>Movement trajectory of the cart during control operation that uses control knowledge obtained after 550 trials.</td>
<td>51</td>
</tr>
<tr>
<td>3.14</td>
<td>The angular dynamics of aerial hovering vehicle. (ArDrone by Parrot)</td>
<td>52</td>
</tr>
<tr>
<td>3.15</td>
<td>Parameters for determining rewards in Learning Control System for rapid position control.</td>
<td>54</td>
</tr>
<tr>
<td>3.16</td>
<td>The structure of the Learning Control System for rapid position control.</td>
<td>55</td>
</tr>
<tr>
<td>3.17</td>
<td>Obstacles and target location assigned in the simulations of Learning Control System for rapid position control.</td>
<td>56</td>
</tr>
<tr>
<td>3.18</td>
<td>Successful control operation for the simulation with assigned target state.</td>
<td>58</td>
</tr>
<tr>
<td>3.19</td>
<td>Successful control operation without obstacles in direct path. (Target State 1)</td>
<td>58</td>
</tr>
<tr>
<td>3.20</td>
<td>Successful control operation with obstacles in direct path. (Target State 6)</td>
<td>59</td>
</tr>
<tr>
<td>3.21</td>
<td>Accumulation of reward during simulations of Learning Control System for rapid position control.</td>
<td>60</td>
</tr>
<tr>
<td>4.1</td>
<td>System structure for application of compound function.</td>
<td>63</td>
</tr>
<tr>
<td>4.2</td>
<td>Method of multiple functions Learning Control by compound control knowledge.</td>
<td>64</td>
</tr>
<tr>
<td>4.3</td>
<td>Reward for control in Goal Attainment Function.</td>
<td>66</td>
</tr>
<tr>
<td>4.4</td>
<td>Reward for control in Obstacle Avoidance Function.</td>
<td>67</td>
</tr>
<tr>
<td>4.5</td>
<td>Structure of Learning Control System by compound function in case of 2 control functions. (Goal and Obstacle)</td>
<td>68</td>
</tr>
<tr>
<td>4.6</td>
<td>Specification of the control device for experiments.</td>
<td>69</td>
</tr>
<tr>
<td>4.7</td>
<td>Field map for simulation of Learning Control System by compound function.</td>
<td>71</td>
</tr>
<tr>
<td>4.8</td>
<td>Training operation for achieving goal using Learning Control.</td>
<td>72</td>
</tr>
<tr>
<td>4.9</td>
<td>Accumulated reward for goal knowledge over simulation episode.</td>
<td>72</td>
</tr>
<tr>
<td>4.10</td>
<td>Training operation for avoiding obstacles using Learning Control.</td>
<td>73</td>
</tr>
<tr>
<td>4.11</td>
<td>Accumulated reward for obstacle knowledge over simulation episode.</td>
<td>73</td>
</tr>
<tr>
<td>4.12</td>
<td>Training results of compound knowledge in simulation.</td>
<td>74</td>
</tr>
<tr>
<td>4.13</td>
<td>Evaluation of real operation with robot.</td>
<td>75</td>
</tr>
<tr>
<td>4.14</td>
<td>Movement results of the evaluation.</td>
<td>75</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Reward settings for assigned state clusters. ........................................... 19
2.2 Parameters description for cart-pendulum device. .............................. 22
2.3 Parameters of the cart-pendulum device. ............................................. 24
2.4 Parameters for Q-learning of Learning Control System by substitute targets. 24
2.5 Initial state and target state of the simulations for Learning Control System by substitute target. .............................................................. 25

3.1 Pre-experimental results for determining the output required by the cart-pendulum device for emulating the angular control of aerial hovering vehicles. 45
3.2 Q-learning parameters for Learning Control System for rapid position control. 47
3.3 Time required to complete a position control during a successful operation. 51
3.4 Specifications of the simulated aerial hovering vehicle. .......................... 55
3.5 Q-learning parameters of Learning Control System for rapid position control. 56

4.1 Specifications of the simulated control device for Learning Control System by compound function. ............................................................ 69
4.2 Parameters for Q-learning in Learning Control System by compound function. 70
Abstract

Human has the ability to learn and decide their action based on experiences when confronting a problem. Human decision often involves multi-functionality, where multiple control functions are applied for achieving a single goal. Conventional control often involves human in providing commands which mostly depends on the human decision. However, these decisions commonly involve single control function where multi-functionality can not be provided without human assistance.

Learning Control helps a machine constructs its own control knowledge autonomously through operation experiences. The development of Control Knowledge through Learning Control would require a period of training that could involve a number of failures among successful attempts. The Control Knowledge obtained is usually limited to single control function based on the training environment with less flexibility in varying environment.

Learning Control Systems with multiple functions could provide a wider range of control options against any environment. In this research, Learning Control System with multi-functionality is designed and developed. Here, application of Learning Control with multi-functionality provides a more human-like control operation with ability to adapt and consider the surrounding environment during control operation. The designs were evaluated through experiments and simulations where results confirm the effectiveness of the designed system. Through these results, the designs of multi-functions Learning Control may provide a safer and reliable control on control devices including complex non-linear control device.
Chapter 1

Introduction

1.1 Research Background

Human perform actions in order to complete task or react to surrounding environment. We render these actions in functions form. The actions are naturally based on purposes, which commonly act as goals. Successes and failures in achieving these goals are recorded in the human mind as knowledge, for references during future attempts. This form of learning represents human intelligence for being self-sustainable that is important in improving our skills for solving surrounding problems.

Applying such intelligence in machines has been an issue surrounding many researchers. Methodologies for self-sustained autonomous machines have been well developed and various new methods and ideas are continuously being proposed in order to reduce human intervention in managing these machines. Providing actions of machines in form of functions help machines to self-evaluate their actions. Human-like functions are one of the focuses of these methods and application may provide methods for self-sustained autonomous machines that could react and adapt to surrounding environment.

1.1.1 Multi-Functionality

Human functions are not limited to individual components where each functions only reacts to a single goal. A goal may require multiple human functions to be obtainable. For example, in case of hurdle race, two human functions of jumping and running are combined to cross the finishing line which acts as a goal. Here, multiple functions are utilized, where a professional with only either jumping or running skills are not certain to be capable of achieving the finishing line perfectly. The above ability here is described as Multi -
functionality. Through Multi-Functionality, an action can be learned and decided by multiple knowledge of skill, and applied when confronting a problem that cannot be solving by a single function. Here, Multi-functionality can be described as a quality of utilizing multiple functions for performing a single goal.

A device with multi-functionality could render an action that considers multiple characteristics in surrounding environment through application of knowledge of skills from various environments. A device with conventional control method only utilizes control command that produces action based on a single function. Method of self-sustained machines could only utilize a single function to become sustainable and lack of flexibility in confronting foreign characteristics simultaneously. Multiple control option is needed in self-sustained machines in order to become autonomous. Multi-Functionality may provide a wide range of control option against any environment in self-sustainable machines.

1.1.1.1 Multi-Functionality against Non-Linearity

Most control method considers linearity in a device for deciding control option. A device with non-linearity will not able to utilize a single control method for the entire system due to parameters that would render the system unstable at a certain state. For example, a pendulum-cart device has two different states that require different control methods for operation. Multiple functions are needed to manage these multiple states. Conventional control method such as Cascade PD Control can only provide two functions for swing and stabilization control. In case of more functions required, such method could not manage to perform successfully.

![Figure 1.1: States for control of Cart Pendulum System.](image)
Non-linearity also exists in our common devices such as vehicles. Non-linear Control in machines is complex and hard without an expert human knowledge in the control system. Aerial hovering vehicles such as helicopters require multiple functions for managing multiple states using the Thrust and Cyclic.

Manipulation of angular orientation with thrust can provide position transition but requires skills in multi-functionality. Human multi-functionality provides expert control of machines with non-linearity. Providing multi-functionality in a non-linear control system could provide a safe and reliable control as good as an expert human.

Human multi-functionality provides expert control of machines with non-linearity due to utilization of multiple knowledge of skill when managing the machines. Through skills of angular orientation and hovering thrust manipulation, expert human pilots are able to perform radical movement of such machines in precision, for example, during position transition of the vehicles. They may react to surrounding environment while still maintaining stability of the machine that is easily affected by unstable states. Therefore, providing quality of multi-functionality as well as human-like functions in a non-linear automatic control system could provide a safe and reliable control replacing an expert human.
1.1.1.2 Embedding Human Knowledge for Multi-Functionality

Embedding human like functions in a system through application of Intelligent Control that provides detailed decision during control operation in a certain environment. Various method concerning intelligent control system may help provides control alternative to an expert skills in controlling a device. Intelligent Control System provides autonomous development of control knowledge together with autonomous development of control strategy on a device. Control Knowledge and Control Strategy are developed depending on a human control decision together with the environment feedback. The developed Control Knowledge and Control Strategy may perform as well as an expert human controlling the machine, reducing the command burden on the human. However, the embedded human functions in the control knowledge are usually constrained to a single function.

Figure 1.4: Control operation during position transition of aerial hovering vehicles.

Figure 1.5: Structure of functions in an Intelligent Control System.
Learning is one of the qualities for developing control knowledge in Intelligent Control System. Control knowledge may be developed through learning method such as trial and error processes. Machine Learning provides option in generating development of control knowledge in an intelligent control system. Control knowledge may be developing through experiences by method in Machine Learning such as Reinforcement Learning. Development of control knowledge helps an Intelligent Control System remain self-sustained and adaptable to changes in surrounding environment. Therefore, new functions may be learned through the learning process giving quality of multi-functionality to the Intelligent Control System.

1.1.2 Learning Control

Learning is generally defined as the process of acquiring new knowledge. The process of acquiring new knowledge needs one to represent the knowledge in some form, as learning is constructing or modifying representations of what is being experienced [4]. The representations meaning varies depending on the knowledge it represents which can be in a form of algorithm, simulation models, control procedures and such.

The term of Machine Learning is derived by the ability of a machine on acquiring knowledge from experiences or a set of data. Mitchell [1] defines learning as performance improvements at some tasks through experience. Mitchell defines it precisely as,

\[
A \text{ computer program is said to learn from experience } E \text{ with respect to some class of tasks } T \text{ and performance measure } P, \text{ if its performance at tasks in } T, \text{ as measured by } P, \text{ improves with experience } E. 
\]

To have a well-defined learning problem, three features concerning class of tasks, the measure of performance to be improved, and the source of experience must be defined. Thus, Machine Learning aims to have a computational mechanism that can learn to improve knowledge through operational experience.

1.1.2.1 Reinforcement Learning

Reinforcement Learning is known as trial and error style learning process that learns to map situations and actions by maximizing a numerical reward signals [4]. All Reinforcement Learning agents may have explicit goals. Using its experience, the agents improve its performance over time. Aspect of their environments can be sense and actions are
changeable to influence their environment. Reinforcement Learning acquires action rules for adapting with the surrounding environment. Reinforcement Learning operates through interactions and acquires knowledge by categorizing actions using rewards, optimizing the best possible action required in order to complete a task [3].

Reinforcement Learning normally consist four main sub-elements in its system [4]. A Policy to determine behaviour, a Reward Function to determine reward, a Value Function to emulate knowledge and sometimes a Model Environment to mimics the property of the environment. The relation between these elements can be seen in Figure 1.6.

A Policy defines the agent behaviour. Policy perceives state mapping of the agent environment to actions to be taken when is those states. A Policy might be a simple function or a lookup table but sometimes involves extensive computation such as search process. Policy is the core component of a Reinforcement Learning agent since it alone determines the behaviour of the agent. A Reward Function defines the goal for the agent. Reward Function maps each perceived state to a single number which known as reward, indicating the desirability of the state. The purpose of Reinforcement Learning is to maximize these rewards in an operation. In other words, Reward Function defines the good and bad of an action for the system to operate. Reward Function is needed to alter the policy. Generally, actions with low reward will less likely to be selected by the policy repeatedly.

While Reward Function indicates good and bad action immediately, a Value Function acts as knowledge of the good and bad action experienced in a long term operation. The Value Function represents the value of states and indicates the desirability of the state reoccurrences in a long term operation. A state may have low rewards but high in value since it is regularly followed by other state that can yield high rewards. Therefore, a

Figure 1.6: Interaction between policy, reward function and value function.
method of converging the value of state-action pairs $Q(s_t, a_t)$ into an average is called the \textit{Q-Learning} algorithm as

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha (r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)).$$ \hspace{1cm} (1.1)$$

Reinforcement Learning may provide the ability of learning for a control system, for example, on a mobile robot. Programming all possible tasks for the robot can be hard and difficult, simplified by applying a learning ability for the mobile robot. In some cases, a control system would have problems with wear and tear in the control object hardware that can cause unprecedented misconfiguration during long term operation. Provided an ability to learn, the control system may adapt to the condition of its control object on any uncertainty and unforeseen changes by continuous self-calibration.

Learning control refers to the process on developing control strategy in a particular control system by trial and error [6]. This is a branch of Reinforcement Learning in control application where agent learns by analysing good and bad influences those results from its own action during control operation. Learning control resembles the way that humans and animals learn to construct their knowledge of movement strategy based on interaction with the environment.

\subsection*{1.1.2.2 Absence of Multi-Functionality in Reinforcement Learning}

Through Learning Control, control knowledge of a control function can be created through the training by Reinforcement Learning. However, conventional Reinforcement Learning method does not provide application of more than one control function within a Learning Control System. Execution of more control function within a Learning Control System would require application of multiple learning processes within a control system. Methods concerning application of learning processes in Learning Control vary depending on application of the control device and the purpose of the system.

Multi-agent Reinforcement Learning is one of the method concerning application of multiple learning process within a Learning Control System. Application of multiple agents in Reinforcement Learning utilizes learning process for multiple agents, where these agents interact between each other in developing the desired control knowledge [49]. State transitions in the case of multi-agent Reinforcement Learning are the results of the joint action that was performed by the agents within the system. Rewards are evaluated through the joint action, and the control knowledge is updated through a joint policy. In this case, the goal
can be determined through adaptation of the dynamic behaviour between these agents [27]. In case of controls, through multi-agent Reinforcement Learning, dynamic behaviour of the agents performs an action that requires the agents to adapt through an environment but the functionality of these agents is limited [28]. Such behaviour, may have exploration task, where the agents has a function of maintaining a group of moving targets within the sensor range [50] [51]. In overall, multi-agent Reinforcement Learning only focuses on application of multiple agents by Reinforcement Learning for utilization of a primary function.

Hierarchical Reinforcement Learning applies a learning process for improving the reliability of Reinforcement Learning application in real-world problem. Conventional Reinforcement learning methods provides solution in providing adaptable control knowledge in form of value functions. However, the bigger the size of state-space variables, the performance of Reinforcement Learning reduces and would require a large scale of computational effort for the update of the control knowledge. Hierarchical Reinforcement Learning accelerates the reliability of the learning process, where state variables are independent from one another, ignoring irrelevant aspects when solving a sub-task [42]. Hierarchical Reinforcement Learning provides a form of decision management in a system, where sub-task will be surveyed by parent task, providing only relevant action depending on the sub-task performance. In this case, value function of the parent task is separated into value functions of sub-task, where learning process occur ignoring irrelevant sub-task during a precise operation. The value functions of sub-task are then converged into performing a value function of parent task [43]. The main purpose of such method is mainly to increase reliability of the learning process in a certain function that requires monitoring of multiple states in a more accelerated pace.

Applying Learning Control System with a number of Control Functions could provide a wider range of control option against any environment. The Control System should be able to develop and apply the required Control Function according to necessity and could provide a more versatile control operation. Current Reinforcement Learning does not emphasize multi-functionality in a control system. Therefore, a method of applying a number of control knowledge with decision management that can provide cooperation between each provided control function is deemed necessary for a quality of multi-functionality in control system.
1.2 Research Objective & Contents

Through multi-functionality and Learning Control, an idea of a control system that is capable of utilizing any functions while being self-sustainable is possible. Researches concerning Multi-functionality in Learning Control are unfolded in this dissertation. This dissertation provides control design that makes use Learning Control in providing multi-functionality in a Control System.

1.2.1 Research Objective

The objective of this research is to design and develop methods of applying Learning Control that provides multiple control function in control command autonomously during a control operation. Through this research, a control system that is self-sustainable, reliable and adaptable to its surrounding environment motivates the development of methods in achieving Multi-functional Learning Control System. Characteristics of such system can be divided into three qualities.

Firstly, the system is believed to be able to provide safe and reliable control operation in any environment through development of the control knowledge according to successes and failure during control attempts. Experience from past control attempts can be referred to while safer future attempts are being planned. Consecutive attempt continues the development of the control knowledge that renders the system upon becoming an expert system with expert control knowledge.

Secondly, the system is believe to be able to reduce dependency on human intervention by self-sustaining system development during control operation in a certain environment. Control Decision can mostly be provided by the system based on the control knowledge developed, reducing the need of human commands. Thus, reduces the requirement on skills on the human operators while maintains the expertise in executing the control operation.

Thirdly, the system is believed to be able to provide decision management in a Learning Control System, that could considers multiple functions during execution. Here, the system may provide wide range of control options during control operation while considering changes in surrounding environment.

The above characteristics provide ideas in designing systems that reflects the motivation of this research. Design of systems that consists above characteristics is unfolded in this dissertation in three chapters.
1.2.2 Research Content

Here, three phases of development were organized to fulfill the objective of applying multi-functionality in a control system. First, for applying multi-functionality in a non-linear system, a Substitute Target based Learning Control System with Multiple Control Function was designed. Secondly, for applying human-like multi-functionality in a control system, Learning Control System with multiple control function by multiple source of control knowledge was designed. Finally, for applying human-like decision management with multi-functionality, Learning Control System with multiple control function by Compound Function was designed.

In this chapter, the background of the research is explained, concerning motivation in application of multi-functionality in controls by Learning Control System. Later, background research concerning Learning Control System is introduced, which emphasizes lack of focuses in application of Learning Control concerning multi-functionality. This leads to the objective of this research which explains the needs and potential of a Learning Control System that emphasizes on multi-functionality.

In chapter 2, a design of Learning Control System with multiple control function that applies substitute target for multi-functionality is introduced and applied on cart-pendulum control system. The designed System focuses on providing multi-functionality in the pendulum swing up control that may consider surrounding constraints for achieving the inverted states. The system applies Learning Control in producing substitute targets for the cart position transition which swings the pendulum simultaneously. The substitute targets act as intermediate targets that help the system consider optimal cart movements to provide swinging motion on the pendulum that propels it towards the inverted states under the influence of environmental constraints.

In chapter 3, a design of Learning Control System with multiple control functions by multiple source of control knowledge is introduced and applies on control systems of cart-pendulum and aerial hovering vehicle. The design focuses in applying multi-functionality through application of multiple source of control knowledge. It was applied on rapid position controls of aerial hovering vehicle that was simulated through cart-pendulum controls. The design was improved for control of aerial hovering vehicle among constraints that was applied on simulation of aerial hovering vehicle. The designed utilizes multiple sources of control knowledge for providing controls of angular orientations on the aerial hovering vehicle.

In chapter 4, a design of Learning Control System with Multiple Control function by
Compound Function is introduced and applies on control system of a mobile robot. The design focuses in applying multi-functionality through application of multiple source of control knowledge that merges through utilization of Compound Function. It was applied for position transition and obstacle avoidance control of the mobile robot that was simulated and later applied on a real world operation. The design utilizes Compound Function for creation of Compound Knowledge that consists of compounded control information from the sources.

Finally, the designs in this research are concluded together with suggestion of further research.
Chapter 2

Multiple-Functions Learning Control by Substitute Target

Designing Learning Control with a quality of multi-functionality requires recognition of continuing states during controls operation. Like human recognizing positions for the next step during walking, the positions of those steps reacts as substitute targets where the main target is the desired location of the human. The substitute targets provide options of manoeuvre, where certain action, in case of walking, can be operated flexibly along constraints during the manoeuvre. Therefore, one of the designs concerning Learning Control with multiple functions involves application of substitute target in the Learning Control System.

2.1 Substitute Target

Conventional Reinforcement Learning involves application of state-action pair for providing control knowledge of a certain control operation. Optimum action is learned based on the states of the control object through the success and failure attempted during the control operation. Comparing such application to human, human decide a target or goal before applying an action. For example, in case of walking, a target for steps is determined before the action of walking is applied. A wrong position would render the walking operation colliding with constraints, or heading in the wrong direction. Targets make configuration easier, since target is a part of state elements, such as steps to location of human. Most controls of actuators apply targets as reference for feedback during control operation as well. Multiple target states provide multiple choices of actions for achieving goal and such supporting target states is defined here by substitute targets.

Substitute target is necessary for flexibility in providing system respond to the change of
situation in environment. Controls by substitute target provide flexible action in which important for having an adaptable Learning Control System for such application on machines with non-linearity. In this case, substitute target provides enhancements of action through continuity in applying those targets. For example as shown in figure 2.1, generating an initial action $a_1$ and continues with action $a_2$ during a control operation provides continuous action, or an action with increasing magnitude. Such function may provide precision of a higher magnitude action and reduces the risk of rampaging actions.

Figure 2.1: Substitute target provides continuity of action by providing an intermediate state.

Substitute targets may also provides rearrangements of control manoeuvre for adapting with constrained environment. During operation in a constrained environment, interference by constraint state would jeopardize the control operation, where rearrangement of controls manoeuvre are necessary. Figure 2.2 referred to a case, where substitute target provides rearrangement of actions, creating more substitute targets that provides a safer manoeuver for the control device. When one of the substitute targets are in a constraint state, a new substitute targets can be arrange to provide alternative for the required action. The arrangement of those substitute targets may vary depending on possible combinations that
Possible Action, a,
State
(a) Constraint states interfere with continuity of action.

Figure 2.2: Substitute target can be rearranged to satisfy the need for successful control manoeuvre along constraint states.

provide the required action for fulfilling the goal.

Utilizing substitute target provides flexibility in producing actions in control operations through Learning Control. Safer and more reliable control operation is possible through the application of substitute target in a Learning Control System.

2.2 Utilization of Substitute Target

Utilization of substitute targets may provides a safe and reliable control option for machines with non-linearity. Here, a control system that utilizes substitute target was designed to provide multi-functionality on machines with non-linearity for safe and reliable control operation. Substitute target was applied on a Learning Control System for cart-pendulum device, shown in figure 2.3. Application of such device requires three basic system functions in the designed system; the control function, learning function and recognition function.

In the Learning Control System designed, control function configures the control output for applying forces to the cart based on the targets instructed either from a policy of reinforcement learning or PD control. Learning function updates the knowledge of substitute
target based on the reaction of the pendulum when applying the force to the cart. The recognition function determines the necessary control action depending on the states of the pendulum and the cart, in order to instruct the next required process. Interaction between each functions provide application of substitute target in controlling the cart for applying swing motion on the pendulum towards the desired goal state.

### 2.2.1 Control Function for Substitute Target System

Control function provides control options for the system to apply on the cart. The controls within the function consists two methods; Swing Control and Stabilization Control. Swing control generates forces for increasing the pendulum swing angle when the pendulum is in downward state. The stabilization control generates forces for decreasing the pendulum swing angle when the pendulum is near to inverted state.

During swing control, control output \( u \) provide forces to move the cart for increasing the swing angle of the pendulum. The cart moves to either right or left based on the pendulum angle \( \theta \) and pendulum angular velocity \( \omega \) for intensifying the pendulum swing, increasing the pendulum angle \( \theta \). The initial state of the pendulum was assigned on the downward position where the pendulum angle \( \theta = \pi [\text{rad}] \) as shown in figure 2.4a. The pendulum angle \( \theta \) will increase as the cart moves consecutively until approaching the inverted state. The Learning Control for substitute target was applied on the pendulum swing control. The swing up control arranges targets for cart movement and apply force \( u \) according to those targets.