Human-Machine Co-Learning Design in Controlling a Double Inverted Pendulum

by

Kehao Zhao

B.S, Beijing University of Posts and Telecommunications, 2016

Submitted to the Graduate Faculty of

Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This thesis was presented

by

Kehao Zhao

It was defended on

July 10, 2019

and approved by

Zhi-Hong Mao, PhD, Professor
Department of Electrical and Computer Engineering

Qing-Ming Wang, PhD, Professor
Department of Mechanical Engineering and Materials Science

Xiayun Zhao, PhD, Assistant Professor
Department of Mechanical Engineering and Materials Science

Thesis Advisor: Zhi-Hong Mao, PhD, Professor
Department of Electrical and Computer Engineering

Co-Advisor: Qing-Ming Wang, PhD, Professor
Department of Mechanical Engineering and Materials Science

Human-Machine Co-Learning Design in Controlling a Double Inverted Pendulum

Kehao Zhao, M.S.

University of Pittsburgh, 2019

Effective human-machine interaction is an essential goal of the design of human-machine systems. This, however, is often constrained by the fundamental limitation of the human neural control and inability of the machine's control system in adapting to the time-varying characteristics of the human operator. It is desirable that the control system of the machine can learn to optimize its performance under the behavior change of the human operator. This thesis is aimed at enhancing the machine's control system with learning capabilities. Specifically, an adaptive control framework is proposed that enables human-machine co-learning through the interaction between the machine and the human operator. A dual inverted pendulum system is introduced as an experimental platform. Simulations are performed to implement the control of the two-joint inverted pendulum using the human-machine co-learning controller. The results are compared with those using a controller without learning ability. The parameters of the two controllers are adjusted to explore the effect of the value changing of each parameter on the control performance. Simulation results indicate the superior performance of the proposed adaptive controller design framework.

Table of Contents

Acknowledgment ..... **x**

1.0 Introduction **1**

2.0 Background **4**

2.1 General Description of an Inverted Pendulum System **4**

2.2 Review of Control System Design **5**

2.3 Review of Human Manual Control **7**

3.0 Controller Development **8**

3.1 Double Inverted Pendulum System Development **8**

3.2 Human-Machine Co-Learning Based Controller Design **19**

4.0 Simulations **24**

4.1 Changing the Initial Feedback Gain K **24**

4.1.1 Feedback Gain K_1 = [45 2 15 2.2] **25**

4.1.2 Feedback Gain K_1 = [64 1.5 23 3.2] **29**

4.1.3 Feedback Gain K_1 = [126 0.5 64 8.2] **33**

4.2 Changing the Weighting Matrix Q and R **37**

4.2.1 Weighting Matrix $Q = diag$ [1 , 1 , 0 , 0], R = 1 ..... **37**

4.2.2 Weighting Matrix $Q = diag$ [10 , 10 , 0 , 0], R = 0.1 ..... **39**

4.3 Changing the Updating Rate α **41**

4.3.1 Updating Rate α = 0.5 ..... **41**

4.4 Time Delay **45**

4.4.1 Time Delay τ = 10 Time Intervals **46**

4.4.2 Time Delay τ = 100 Time Intervals 48

4.5 Simulation Results 50

5.0 Conclusion and Future Work 55

5.1 Conclusion 55

5.2 Future Work 55

Appendix A Matlab Code 57

Bibliography 60

List of Tables

Table 1. Parameters of DIP system 9

Table 2. System performance when changing the initial feedback gain K 50

Table 3. System performance when changing the weighting matrix Q and R 51

Table 4. System performance when changing the updating rate α 52

Table 5. System performance with time delay 53

List of Figures

Figure 1. Double inverted pendulum system. 9

Figure 2. Angle between inverted pendulums and vertical direction. 25

Figure 3. Torque applied on the pendulums. 26

Figure 4. Angle between inverted pendulums and vertical direction. 27

Figure 5. Torque applied on the pendulums. 28

Figure 6. Angle between inverted pendulums and vertical direction. 29

Figure 7. Torque applied on the pendulums. 30

Figure 8. Angle between inverted pendulums and vertical direction. 31

Figure 9. Torque applied on the pendulums. 32

Figure 10. Angle between inverted pendulums and vertical direction. 33

Figure 11. Torque applied on the pendulums. 34

Figure 12. Angle between inverted pendulums and vertical direction. 35

Figure 13. Torque applied on the pendulums. 36

Figure 14. Angle between inverted pendulums and vertical direction. 37

Figure 15. Torque applied on the pendulums. 38

Figure 16. Angle between inverted pendulums and vertical direction. 39

Figure 17. Torque applied on the pendulums. 40

Figure 18. Angle between inverted pendulums and vertical direction. 41

Figure 19. Torque applied on the pendulums. 42

Figure 20. Angle between inverted pendulums and vertical direction. 43

Figure 21. Torque applied on the pendulums. 44

Figure 22. Angle between inverted pendulums and vertical direction. 46

Figure 23. Torque applied on the pendulums. 47

Figure 24. Angle between inverted pendulums and vertical direction. 48

Figure 25. Torque applied on the pendulums. 49

Acknowledgment

I would like to express my sincere gratitude to my research advisor, Prof. Zhi-Hong Mao for his patience and considerable help throughout my study and research. Without his consistent and illuminating instruction, this thesis could not have reached its present form.

I would thank Prof. Qing-Ming Wang for introducing me to Prof. Mao, and I really appreciate all the suggestions and help from Prof. Wang throughout my master program.

I would like to thank my committee members: Prof. Qing-Ming Wang, Prof. Zhi-Hong Mao and Prof. Xiayun Zhao for their valuable time.

I would have my gratitude to the lab members in Dr. Sharma's research group: Xuefeng Bao, Ziyue Sun and Qiang Zhang for their generous help.

Finally, I would thank my family and friends for their supports and encouragements.

1.0 Introduction

Human–machine system is a system that integrates the functions of a human operator and a machine [1]. The study of human-machine systems aims to develop the theories and methodology to analyze, design and evaluate the interaction between human and machine, and the expectation of the system is to combine the advantages of humans and machines. Human have stronger perception and flexible decision-making ability while machine can reduce the probability of occurrence of danger, effectively resist fatigue, and is suitable for repetitive tasks with high precision requirements. Human-machine systems have a wide range of applications in medical, automotive, aerospace and other fields. The aircraft is a typical example. In order to improve the stability of the aircraft and improve the damping characteristics of the aircraft, the manual control system is combined with the automatic control to introduce the stability augmentation system into the manual control system.

As a typical human-machine system, the inverted pendulum under manual control has become a major research direction [2]. The performance of human manual control was evaluated when dealing with a difficult task to understand human movement behavior as well as human limitations. Thus, human strategies in balancing an inverted pendulum with different time delay was studied. An experimental platform was developed consisting of a joystick connected to the computer and a virtual inverted pendulum. The operator balances the virtual inverted pendulum with the joystick and obtains visual feedback through the computer screen. Experimental results showed that the human response is intermittent rather than continuous when balancing a short-length inverted pendulum [3].

However, many difficulties need to be solved in the design of a controller involving human. On one hand, the machine not only needs to adjust the control strategy according to different individuals, but also need to adjust the control parameters online according to the change of human's behavior in the process of human-machine control. For example, in the case of self-driving, vehicles could develop the best control strategy to assist the driver by monitoring and learning the driver's habits. For of exoskeleton systems, the systems need to provide assistance according to the state and the force provided by the user's limbs [4].

On the other hand, it is not well studied how humans adapt to a nonstationary robot whose behavior will change related to human actions. In the early days, humans were abstracted as a single-input linear time-invariant controller. A quasi-linear model was established for analyzing state regulation tasks [5]. In 1970s, the optimal control and estimation theory are applied to the linear dynamical system with white noise to solve the problem of human manual control. It was supposed that after being trained, people could behave optimally under certain circumstances. Humans were abstracted into a feedback controller with the consideration of human inherent psychophysical limitations. Then, a mathematical model of human was established which conclude a reaction time delay, an equalization network and some equivalent neuromuscular dynamics [5], [6]. The model was originally used to simulate a simple single-axis control task. After that, the optimal control model of human was used to analyze more complex manual control tasks involving the control of longitudinal position of a hovering Vertical Take-Off and Landing aircraft [6]. The experiments result showed that the model can reproduce the essential characteristics of humans as well as system performance scores. However, there is no general established rule to ensure a synergistic partnership so more research is needed to ensure the stability and optimization of the human-machine systems.

In this thesis, we concentrate on the design of a human machine co-learning optimal controller. A second-order inverted pendulum system is introduced as an experiment platform. Human controls the lower inverted pendulum and machine controls the upper one. It is supposed that human has the ability to make corresponding changes which are based on the optimization strategy given by the machine to optimize their control parameters. A co-learning controller and a controller without learning ability were established. Simulations were made to control the second-order inverted pendulum, and the two sets of results were compared. At the same time, the initial control feedback gains, iterative updating rate and the weighting matrix Q and R were adjusted in the simulation to get the performances of the co-learning controller.

The thesis is organized as follows: Chapter 1 is a brief introduction to the inverted pendulum system and the control theory applied on them. The innovation of the research is also expressed in this chapter. The background information on inverted pendulum system, control methods, and human manual control are discussed in the second chapter. Chapter 3 shows the development of the second-order inverted pendulum system and the design of the human-machine co-learning controller. The simulation results are shown in Chapter 4. The last chapter summarizes the thesis and shows some ideas for the future work.

2.0 Background

2.1 General Description of an Inverted Pendulum System

In the late 1960s, as a typical unstable nonlinear example, the concept of inverted pendulum was proposed. In 1966, Schaefer and Cannon applied Bang-Bang control theory to stabilize a crankshaft in an inverted position. Since the inverted pendulum system has become the experimental and teaching tool of the automatic control field, researchers have studied both the theoretical research and the experimental research on the control of the inverted pendulum system. In 1969, Lindorff achieved the control of the single inverted pendulum system successfully [7]. In 1972, Sturgeon and Loscutoff controlled the double inverted pendulum [8]. As early as 1970, Bryon and Luenberger first pointed out the application of the observer to reconstruct the state of the system which could realize stable control of inverted pendulum system. In 1985, K. Furutat realized the stable control of the three-stage inverted pendulum [9]. In 1986, Chung systematically identified the inverted pendulum system and designed PD feedback control [9]. In 1989, Anderson successfully generated an optimized feedback controller using function minimization and LyaPunov stabilization [9]. Li controlled the quadruple inverted pendulum successfully in 2002 [10].

2.2 Review of Control System Design

Many papers have implemented the inverted pendulum cart dynamical system with various control schemes.

1. State feedback control

   The process of state feedback control is to multiply each state variable of the system by the corresponding feedback gain and added to the reference input. The sum of the signal is used as the control input of the controlled system. Based on the dynamic model of the inverted pendulum, state equations and output equations are derived using state space theory, and state feedback is applied to control the inverted pendulum. Common methods for using state feedback are: 1) linear quadratic optimal control; 2) pole placement; 3) state feedback control; 4) robust control.

2. PID control

   The PID control is composed of a proportional unit P, an integral unit I, and a differential unit D. Based on the dynamic model of inverted pendulum, the nonlinear model is derived by using state space theory, then the state equation and output equation of the inverted pendulum system are obtained by linearization at the equilibrium point. The PID controller is designed according to the state equation and output equation of the inverted pendulum system to realize the control of the inverted pendulum [11], [12].

3. Adaptive Control

   The object of adaptive control is a system with a certain degree of uncertainty. The uncertainty inside the system comes from the unknown structure and parameters of the mathematical model of the controlled object. Uncertain externalities come from unpredictable disturbances and measurement errors. The control parameters are

automatically adjusted as conditions change by methods such as model reference adaptive control or self-tuning control which could optimize the performance of the system. [13].

4. MPC control

   Model predictive control can predict the performance of the system over a limited time frame using mathematical models. The control signal over a limited time range is calculated by minimizing a predefined cost function. The current state is used as an initial condition for the next iteration. Only the first input of the optimal control sequence is applied on the system, and the state is updated in the next iteration. The prediction horizon is shifted one time step backward, and the procedure is repeated [14], [15].

5. Neural network control

   Neural networks are used to handle systems that are difficult to describe with models or rules. Neural network adopts parallel distributed information processing, which is highly fault-tolerant and can process a large number of different types of inputs simultaneously. The goal of the neural network controller is to find an effective way to modify the network connection weight matrix or network structure, so that the control signal output by the network controller can ensure that the output of the system can follow the desired output of the system.[16], [17].

6. Fuzzy control.

   Fuzzy control theory is a kind of advanced control strategy based on fuzzy mathematics and decision making by fuzzy reasoning. Fuzzy control does not have to establish a complete mathematical model for the controlled object, thus simplifying the complexity

of the system design. At the same time, the fuzzy control has strong robustness and better fault tolerance for system parameter changes. [18], [19].

2.3 Review of Human Manual Control

The study of manual control of dynamic systems began in the 1940s and was proposed by Phillips [20], Wiener [21] and Tustin [22]. By 1965, Russell [23], Elkind [24], McRuer and Krendel [25] and McRuer, Graham, Krendel and Reisener [26] abstracted the manual control model into a series of quasi-linear models that could predict human behavior.

In 1960, researchers began to analyze more complex manual control systems. Manual control systems were attempted to abstract into multivariate linear models. Early research relied mainly on classical control theory to extend a single variable to a multivariate case. Later research relies on modern control theory and optimization theory which deal with multivariable systems under the of state space system. In 1970, D. L. Kleinman, S. Baron and W. H. Levison applied optimization and estimation theory to the manual control and established a mathematical model of the human as a feedback controller. In this model, the manual controller is decomposed into a time delay, an equalization network and some neuromotor dynamics [26].

3.0 Controller Development

3.1 Double Inverted Pendulum System Development

The inverted pendulum is a multivariable, unstable and nonlinear system [16] that covers many key features of the control domain such as stabilization, robustness and tracking [27]. The inverted pendulum system is widely used to test the correctness and practicability of the new control theory or method [28]. The inverted pendulum system is cheap and easy to operate which has been widely used as experimental platform. Meanwhile, the inverted pendulum has a wide range of applications in military, aerospace, robotics and general industrial processes, such as robotic walking, rocket launching, and satellite flight [29].

The double inverted pendulum system contains two pendulums, motors and encoders. The two pendulums are joined by a motor and the starting point of lower pendulum is attached to a motor fixed on the ground. Encoders are used to measure the position and velocity of the two pendulums. The double inverted pendulum system is shown in Figure 1 and the parameters of the system are given in Table 1.

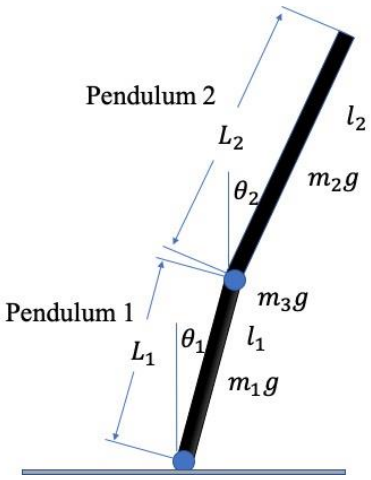

Figure 1. Double inverted pendulum system.

Table 1. Parameters of DIP system

M_1=1.5 Kg, M_2=0.5 Kg, M_3=0.75 Kg	Mass of the lower pendulum, upper pendulum and joint
L_1=1 m, L_2=1.5 m	Length of lower pendulum (= $2l_1$), length of upper pendulum (= $2l_2$)
g=9.8 m/s^2	Acceleration due to gravity
θ_1, θ_2	The angle between pendulum 1, 2 with vertical direction in radian
T	Torque generated by the motor

Dynamic equations of inverted pendulum system is established by the Lagrange equation which can be expressed as

$$\frac{d}{dt}\frac{dH}{d\dot{q}_i} - \frac{dH}{dq_i} = F_i \tag{3.1}$$

where $H = G - V$ is a Lagrangian, F_i is a force vector applied on the generalized coordinates q_i, V and G are potential and kinetic energy of the system respectively.

Kinetic energy of pendulum A is

$$K_A = \frac{1}{6}m_1\left(L_1\dot{\theta}_1\right)^2 = \frac{2}{3}m_1{l_1}^2{\dot{\theta}_1}^2. \tag{3.2}$$

Kinetic energy of motor M is

$$K_M = \frac{1}{2}m_3\left(L_1\dot{\theta}_1\right)^2 = 2m_3{l_1}^2{\dot{\theta}_1}^2. \tag{3.3}$$

Speed at both ends of pendulum B is

$$V_M = L_1\dot{\theta}_1\cos\theta_1\bar{\imath} + L_1\dot{\theta}_1\sin\theta_1\bar{\jmath} \tag{3.4}$$

$$V_N = \left(L_1\dot{\theta}_1\cos\theta_1 + L_2\dot{\theta}_2\cos\theta_2\right)\bar{\imath} + \left(L_1\dot{\theta}_1\sin\theta_1 + L_2\dot{\theta}_2\sin\theta_2\right)\bar{\jmath}. \tag{3.5}$$

Kinetic energy of pendulum B is

$$K_B = \frac{1}{6} m_2 \left(\bar{V}_M{}^2 + \bar{V}_N{}^2 + \bar{V}_M \cdot \bar{V}_N \right)$$

$$= \tfrac{1}{6} m_2 \left(3 l_1{}^2 \dot{\theta}_1{}^2 + l_2{}^2 \dot{\theta}_2{}^2 + 3 l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2 \right)$$

$$= 2 m_2 l_1{}^2 \dot{\theta}_1{}^2 + \tfrac{2}{3} m_2 l_2{}^2 \dot{\theta}_2{}^2 + 2 m_2 l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2. \qquad (3.6)$$

Total kinetic energy of the system is

$$G = K_A + K_M + K_B$$

$$= \frac{2}{3} m_1 l_1{}^2 \dot{\theta}_1{}^2 + 2 m_3 l_1{}^2 \dot{\theta}_1{}^2 + 2 m_2 l_1{}^2 \dot{\theta}_1{}^2$$

$$+ \frac{2}{3} m_2 l_2{}^2 \dot{\theta}_2{}^2 + 2 m_2 l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2. \qquad (3.7)$$

Potential energy of pendulum A is

$$P_A = m_1 g l_1 \cos\theta_1. \qquad (3.8)$$

Potential energy of motor M is

$$P_M = 2 m_3 g l_1 \cos\theta_1. \qquad (3.9)$$

Potential energy of pendulum B is

$$P_B = m_2 g (2 l_1 \cos\theta_1 + l_2 \cos\theta_2). \qquad (3.10)$$

Total potential energy of the system is

$$V = P_A + P_M + P_B = m_1 g l_1 \cos\theta_1 + 2m_3 g l_1 \cos\theta_1 + m_2 g(2l_1 \cos\theta_1 + l_2 \cos\theta_2). \quad (3.11)$$

Therefore, Lagrangian H is

$$H = G - V = \left(\frac{2}{3} m_1 {l_1}^2 + 2m_3 {l_1}^2 + 2m_2 {l_1}^2\right) {\dot{\theta}_1}^2$$

$$+ \frac{2}{3} m_2 {l_2}^2 {\dot{\theta}_2}^2 + 2m_2 l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 - m_1 g l_1 \cos\theta_1$$

$$-2m_3 g l_1 \cos\theta_1 - m_2 g(2l_1 \cos\theta_1 + l_2 \cos\theta_2). \quad (3.12)$$

The derivative for Lagrangian H of generalized coordinate θ_1 is shown as

$$\frac{\partial H}{\partial \theta_1} = -2m_2 l_1 l_2 \sin(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + m_1 g l_1 \sin\theta_1$$

$$+2m_3 g l_1 \sin\theta_1 + 2m_2 g l_1 \sin\theta_1. \quad (3.13)$$

The derivative for Lagrangian H of generalized coordinate θ_2 is shown as

$$\frac{\partial H}{\partial \theta_2} = 2m_2 l_1 l_2 \sin(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + m_2 g l_2 \sin\theta_2. \quad (3.14)$$

The derivative for Lagrangian H of generalized coordinate $\dot{\theta}_1$ is shown as

$$\frac{\partial H}{\partial \dot{\theta}_1} = 2\left(\frac{2}{3} m_1 {l_1}^2 + 2m_3 {l_1}^2 + 2m_2 {l_1}^2\right)\dot{\theta}_1 + 2m_2 l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_2. \quad (3.15)$$

The derivative for Lagrangian H of generalized coordinate $\dot{\theta}_2$ is shown as

$$\frac{\partial H}{\partial \dot{\theta}_2} = \frac{4}{3} m_2 {l_2}^2 \dot{\theta}_2 + 2 m_2 l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1. \tag{3.16}$$

Then, it can be obtained that

$$\frac{d}{dt}\frac{\partial H}{\partial \dot{\theta}_1} = \left(\frac{4}{3} m_1 {l_1}^2 + 4 m_3 {l_1}^2 + 4 m_2 {l_1}^2\right) \ddot{\theta}_1 + 2 m_2 l_1 l_2 \cos(\theta_1 - \theta_2) \ddot{\theta}_2$$

$$-2 m_2 l_1 l_2 \sin(\theta_1 - \theta_2) \dot{\theta}_2 \left(\dot{\theta}_1 - \dot{\theta}_2\right) \tag{3.17}$$

$$\frac{d}{dt}\frac{\partial H}{\partial \dot{\theta}_2} = 2 m_2 l_1 l_2 \cos(\theta_1 - \theta_2) \ddot{\theta}_1 + \frac{4}{3} m_2 {l_2}^2 \ddot{\theta}_2 - 2 m_2 l_1 l_2 \sin(\theta_1 - \theta_2) \dot{\theta}_1 \left(\dot{\theta}_1 - \dot{\theta}_2\right). \tag{3.18}$$

Substituting Lagrangian H in Equation 3.1 and replacing the generalized coordinate with θ_1 and θ_2.

The Lagrange equation can be expressed as

$$\frac{d}{dt}\frac{\partial H}{\partial \dot{\theta}_1} - \frac{\partial H}{\partial \theta_1} = M_1 \tag{3.19}$$

$$\frac{d}{dt}\frac{\partial H}{\partial \dot{\theta}_2} - \frac{\partial H}{\partial \theta_2} = M_2. \tag{3.20}$$

More specifically

$$\left(\frac{4}{3}m_1{l_1}^2 + 4m_3{l_1}^2 + 4m_2{l_1}^2\right)\ddot{\theta}_1 + 2m_2l_1l_2\cos(\theta_1 - \theta_2)\ddot{\theta}_2$$

$$-2m_2l_1l_2\sin(\theta_1 - \theta_2)\dot{\theta}_2\left(\dot{\theta}_1 - \dot{\theta}_2\right) + 2m_2l_1l_2\sin(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2$$

$$-m_1gl_1\sin\theta_1 - 2m_3gl_1\sin\theta_1 - 2m_2gl_1\sin\theta_1 = M_1 \quad (3.21)$$

$$2m_2l_1l_2\cos(\theta_1 - \theta_2)\ddot{\theta}_1 + \frac{4}{3}m_2{l_2}^2\ddot{\theta}_2 - 2m_2l_1l_2\sin(\theta_1 - \theta_2)\dot{\theta}_1\left(\dot{\theta}_1 - \dot{\theta}_2\right)$$

$$-2m_2l_1l_2\sin(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 - m_2gl_2\sin\theta_2 = M_2 \quad (3.22)$$

After simplification, it can be obtained that

$$\left(\frac{4}{3}m_1{l_1}^2 + 4m_3{l_1}^2 + 4m_2{l_1}^2\right)\ddot{\theta}_1 + 2m_2l_1l_2\cos(\theta_1 - \theta_2)\ddot{\theta}_2$$

$$+2m_2l_1l_2\sin(\theta_1 - \theta_2){\dot{\theta}_2}^2 - m_1gl_1\sin\theta_1 - 2m_3gl_1\sin\theta_1 - 2m_2gl_1\sin\theta_1 = M_1 \quad (3.23)$$

$$2m_2l_1l_2\cos(\theta_1 - \theta_2)\ddot{\theta}_1 + \frac{4}{3}m_2{l_2}^2\ddot{\theta}_2 - 2m_2l_1l_2\sin(\theta_1 - \theta_2){\dot{\theta}_1}^2 - m_2gl_2\sin\theta_2 = M_2. \quad (3.24)$$

The operating point of the system is at angles (θ_1 , θ_2) = (0, 0) i.e. when both the pendulum angles are zero. A small perturbation is introduced around operating point and expanded using Taylor series.

The Taylor series with a small angle can be shown as

$$\cos(\theta_1 - \theta_2) \cong 1 \qquad \sin(\theta_1 - \theta_2) \cong \theta_1 - \theta_2 \cong 0$$

$$\cos\theta_1 \cong \cos\theta_2 \cong 1 \qquad \sin\theta_1 \cong \theta_1 \, \sin\theta_2 \cong \theta_2. \tag{3.25}$$

Linearization can be finished at unstable operating point. Then the Lagrange equation can be simplified as

$$\left(\frac{4}{3}m_1{l_1}^2 + 4m_3{l_1}^2 + 4m_2{l_1}^2\right)\ddot{\theta}_1 + 2m_2l_1l_2\ddot{\theta}_2$$

$$-m_1gl_1\theta_1 - 2m_3gl_1\theta_1 - 2m_2gl_1\theta_1 = M_1 \tag{3.26}$$

$$2m_2l_1l_2\ddot{\theta}_1 + \frac{4}{3}m_2{l_2}^2\ddot{\theta}_2 - 2m_2l_1l_2(\theta_1 - \theta_2){\dot{\theta}_1}^2 - m_2gl_2\theta_2 = M_2. \tag{3.27}$$

The Lagrange equation can be expressed in matrix form

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a_1 & a_2 \\ 0 & 0 & a_3 & a_4 \end{bmatrix}\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ f_1 & 0 & 0 & 0 \\ 0 & f_2 & 0 & 0 \end{bmatrix}\begin{bmatrix} \theta_1 \\ \theta_2 \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \tag{3.28}$$

where

$$a_1 = \frac{4}{3}m_1{l_1}^2 + 4m_3{l_1}^2 + 4m_2{l_1}^2$$

$$a_2 = 2m_2l_1l_2$$

$$a_3 = 2m_2l_1l_2$$

$$a_4 = \frac{4}{3} m_2 {l_2}^2$$

$$f_1 = m_1 g l_1 + 2m_3 g l_1 + 2m_2 g l_1$$

$$f_2 = m_2 g l_2.$$

Let

$$a = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a_1 & a_2 \\ 0 & 0 & a_3 & a_4 \end{bmatrix}$$

$$b = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ f_1 & 0 & 0 & 0 \\ 0 & f_2 & 0 & 0 \end{bmatrix}$$

$$c = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Multiply both sides of the equation by a^{-1}

$$A = a^{-1} b$$

$$B = a^{-1} c.$$

It can be obtained that

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 14.25 & -2.67 & 0 & 0 \\ -14.25 & 12.47 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.73 & -0.73 \\ -0.73 & 3.39 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Let

$$x = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix}$$

$$\dot{x} = \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix}.$$

The state-space representation is shown as

$$\dot{x} = Ax + Bu \tag{3.29}$$

$$y = Cx. \tag{3.30}$$

Or explicitly:

$$\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 14.25 & -2.67 & 0 & 0 \\ -14.25 & 12.47 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.73 & -0.73 \\ -0.73 & 3.39 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \tag{3.31}$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix}. \tag{3.32}$$

3.2 Human-Machine Co-Learning Based Controller Design

The main purpose of the controller is to make the two pendulums reach equilibrium positions without large oscillations and excessive angles and speeds under the effect of random disturbances. It can be achieved by applying suitable torque generated by the motor at the joint. The controllers designed for the double inverted pendulum system are shown below.

Linear quadratic regulator (LQR) is a kind of optimal control techniques. The LQR optimal design refers to the designed state feedback controller K to make the quadratic performance index function J take the minimum value and the optimal control input can be obtained by solving the algebraic Riccatti equation[30]. In this paper, the state feedback controller is designed using the linear quadratic regulator.

The nonlinear system equations are linearized at the unstable equilibrium point which have the initial conditions

$$x_0 = [0 \quad 0 \quad 0 \quad 0]^T$$

the linearized state-space equation is shown in equation 3.28.

The feedback control can be expressed as

$$u(t) = -Kx(t). \tag{3.33}$$

The performance index function is shown as

$$J = \frac{1}{2}\int_0^\infty [x^T(t)Qx(t) + u^T(t)Ru(t)]dt \tag{3.34}$$

where Q is positive-semi definite matrix, R is positive definite matrix.

The performance index function should be minimized to obtain the optimal control strategy.

The input u is determined by solving the Riccati equation which can be shown as

$$P(t)A + A^T P(t) - P(t)BR^{-1}(t)B^T P(t) + Q(t) = 0 \quad (3.35)$$

where *P(t)* is the Riccati matrix. The optimal control input $u^*$for any initial state x(0) is given as

$$u^* = -K^* x(t) = -R^{-1}B^T P(t)x(t). \quad (3.36)$$

K is the linear optimal feedback gain matrix which can be shown as $K = -R^{-1}B^T P(t)$.

The state weighting matrix *Q* and input weighting matrix *R* can significantly influences the performance of the system. There is no straightforward way to obtain the weight matrixes. Usually, the weight matrix *Q* and *R* that meets the design requirements is found by repeated simulations and experiments. The first term of the performance index function $\frac{1}{2}x^T(t)Qx(t)$ can be used to measure the error between the given state and the actual state of the system during the entire control period. The integral of $\frac{1}{2}u^T(t)Ru(t)$ represents the energy consumed during the control process and can be used to measure the amount of energy consumed. When the value of an element in the *Q* matrix increases, the dynamic performance of the corresponding x increases. At the same time, oscillations are created. When the elements in the *R* matrix increase, the magnitude of the controlled states decreases which means that the energy consumption of the system decreases as *R* increases[31].

For the DIP system, the quadratic performance index of LQR controller should keep the inverted pendulum within the linear range. Meanwhile, control of the upper inverted pendulum takes precedence over the control of the lower inverted pendulum.

As mentioned above, assume that the people involved in the experiment are well trained. So the human control strategy can be abstracted into a LQR controller. Without loss of generality, Q and R are chosen as $Q = diag[10,10,0,0]$, $R = [1]$.

Assume that the machine has learning ability. As the human control strategy changes, the model of the lower inverted pendulum will change accordingly. The machine takes the changed model as a new system, and updates the optimal feedback gain k according to the LQR controller. Q and R are chosen as $Q = diag[10,10,0,0]$, $R = [1]$.

$$\begin{aligned} \dot{x} &= Ax + Bu \\ &= Ax + [B_1 \quad B_2]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= Ax + B_1u_1 + B_2u_2. \end{aligned} \tag{3.37}$$

Assume that human have the control strategy K_1

$$u_1 = -K_1x. \tag{3.38}$$

Substitute $u_1$into the state-space equation

$$\begin{aligned} \dot{x} &= (A - B_1K_1)x + B_2u_2 \\ &= A_2x + B_2u_2 \end{aligned} \tag{3.39}$$

where

$$A_2 = A - B_1K_1.$$

By applying the LQR controller, the optimal feedback gain of the machine is obtained which can be shown as K_2.

Similarly, the update of the machine control strategy will result in changes in the parameters of the upper inverted pendulum. Experimental participant also make this change and

the original model as a whole and updates the optimal feedback gain K_1 according to the LQR controller. *Q* and *R* matrix remain unchanged.

The control input

$$u_2 = -K_2 x. \tag{3.40}$$

Substitute $u_2$into the state-space equation

$$\begin{aligned} \dot{x} &= (A - B_2 K_2)x + B_1 u_1 \\ &= A_1 x + B_1 u_1. \end{aligned} \tag{3.41}$$

where

$$A_1 = A - B_2 K_2.$$

The A matrix of the state space is updated, and the human optimal feedback gain K_1 is updated correspondingly.

Repeat this process, using $K_1 x$ and $K_2 x$ as part of the state space and get the corresponding optimal solution ${K_1}^*$ and ${K_2}^*$ at this time interval. The optimal feedback gain vector of human and machine are obtained through iterative calculation.

The process of learning between human and machine is transformed into the process of iteratively calculating the optimal feedback gain vector. In this process, it is assumed that the learning speed of the human is slower than that of the machine . For example, the human control strategy K_1 is updated in two time intervals and the machine's control strategy K_2 is updated in one time interval.

Meanwhile, Supposed there is a delay for the experimenter in applying the learned control strategy to the actual operation while the machine can apply the updated strategy to the input in real time. The time delay is three time intervals. In the first three time intervals, the control strategy applied by human is the initial value. Starting from the fourth time interval,

assuming that the calculated control strategy is $K_1(n)$, the actual applied control strategy is $K_1(n-3)$.

To prevent instability due to excessive changes of feedback gain K, a updating rate α is proposed. Rather than applying the optimal feedback gain K^* calculated by LQR to the system directly, the new updating law is shown as

$$K_{new} = K_{old} + \alpha(K^* - K_{old}) \tag{3.42}$$

where K^* is the optimal feedback gain calculated by LQR, K is the feedback gain applying on the system.

4.0 Simulations

In order to test the performance of the LQR controller designed in chapter 3, simulations were conducted to compare the performance of the controller under different control parameters. In the following section, the procedure and the results of the simulations are presented.

In the MATLAB simulation, a state space model was built for testing the performance of the controller. The time interval was chosen as 5ms. The MATLAB code for the controller is shown in Appendix A.

The simulation results including the angles of two inverted pendulums, and control input.

4.1 Changing the Initial Feedback Gain *K*

Assuming the state weighting matrix *Q*, input weighting matrix *R* and the update rate α are constants. Comparing the simulation results by using MATLAB when the initial feedback gain *K* changed.

In this section, the Co-learning model and the model in which machine does not have the ability to learn were simulated. The initial condition is that the lower inverted pendulum is vertical, and the upper inverted pendulum has an initial angle of 5 degrees. When the system reaches the steady state under the control of the controller described above, the number of time intervals required for the upper and lower inverted pendulum to converge below $5 \cdot 10^{-5}$

radians, the maximum deviation angle of the lower inverted pendulum, the maximum output torque of the two motors and the minimum output torque of the two motors are compared.

In the simulation below, let $Q=diag[10 , 10 , 0 , 0]$, $R=1$, updating rate $\alpha = 0.1$;

4.1.1 Feedback Gain K_1 = [45 2 15 2.2]

1 Co-learning model

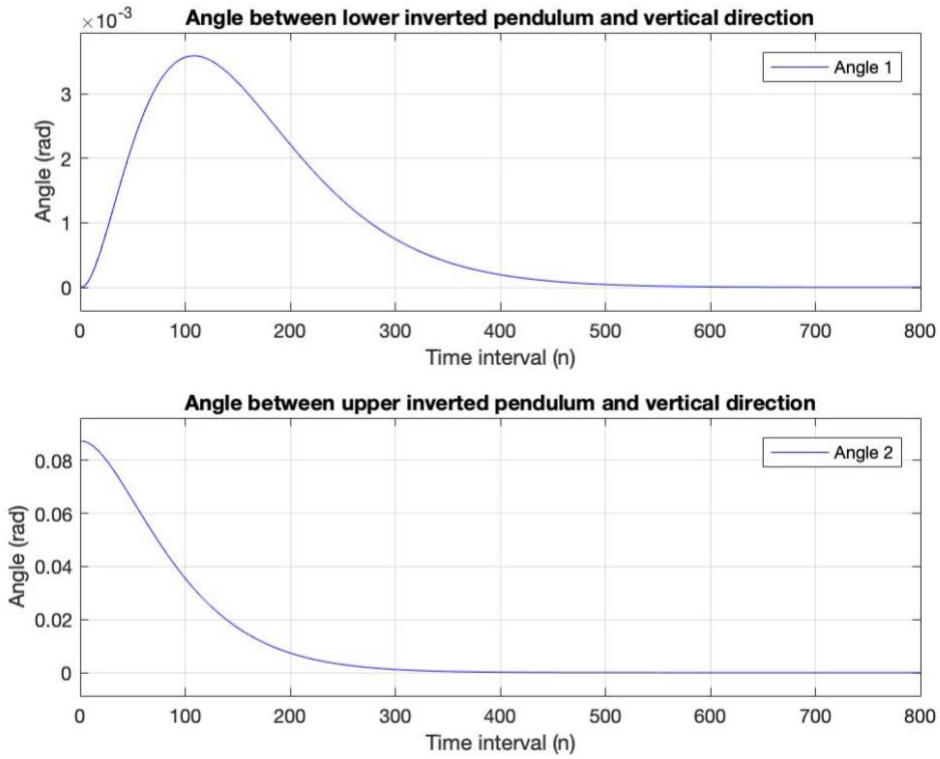

Figure 2. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0036 Rad at time n = 108 time intervals.

It takes 492 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 458 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

Where $\angle 1$ denote the angle between lower inverted pendulum and vertical direction and $\angle 2$ denote the angle between upper inverted pendulum and vertical direction.

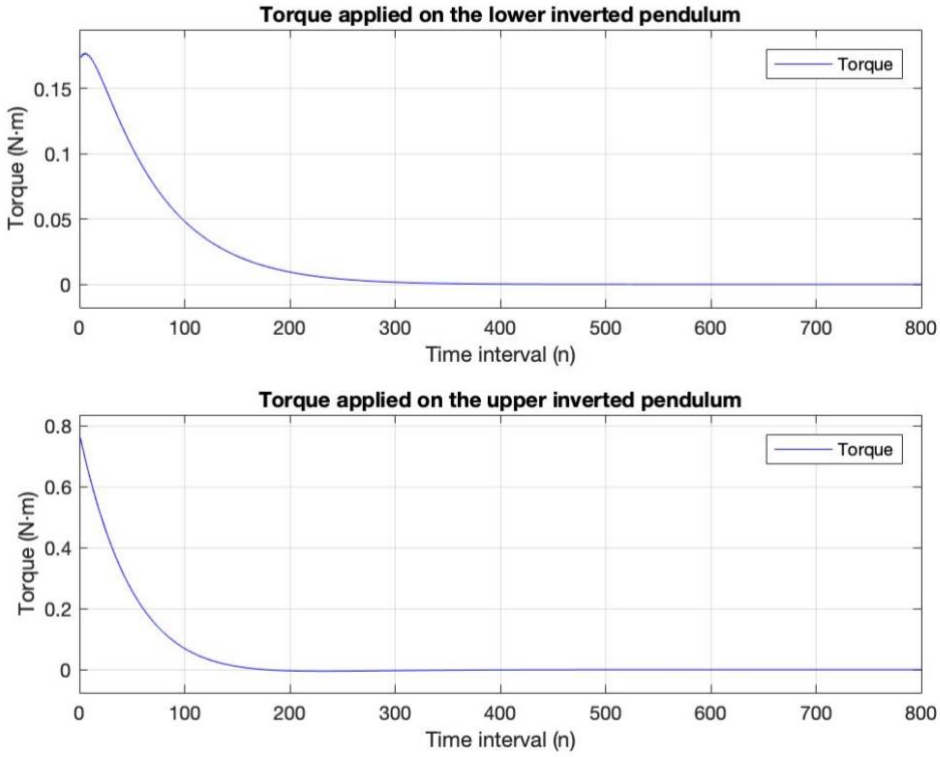

Figure 3. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1773 N·m.

The maximum torque of the upper motor is 0.7603 N·.

2 No learning ability model

The initial control strategy of human K_1=[45 2 15 2.2]. The corresponding machine control strategy can be calculated by Equation 4.6 - 4.7.

Which is K_2=[2.38 8.72 2.08 2.59].

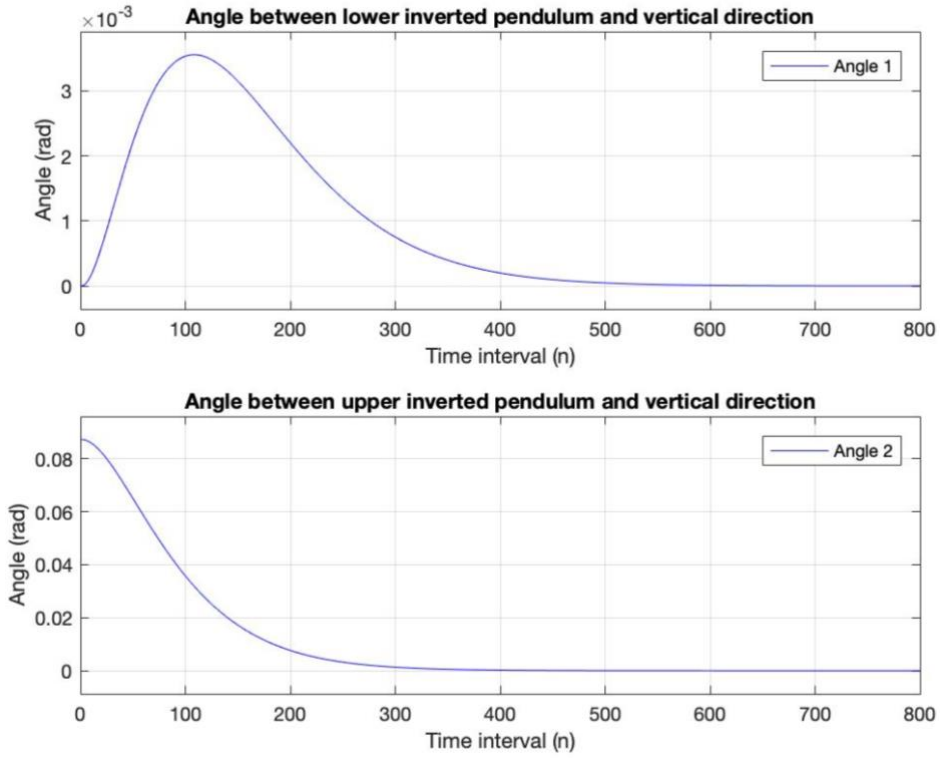

Figure 4. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0036 Rad at time n = 108 time intervals.

It takes 496 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 483 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

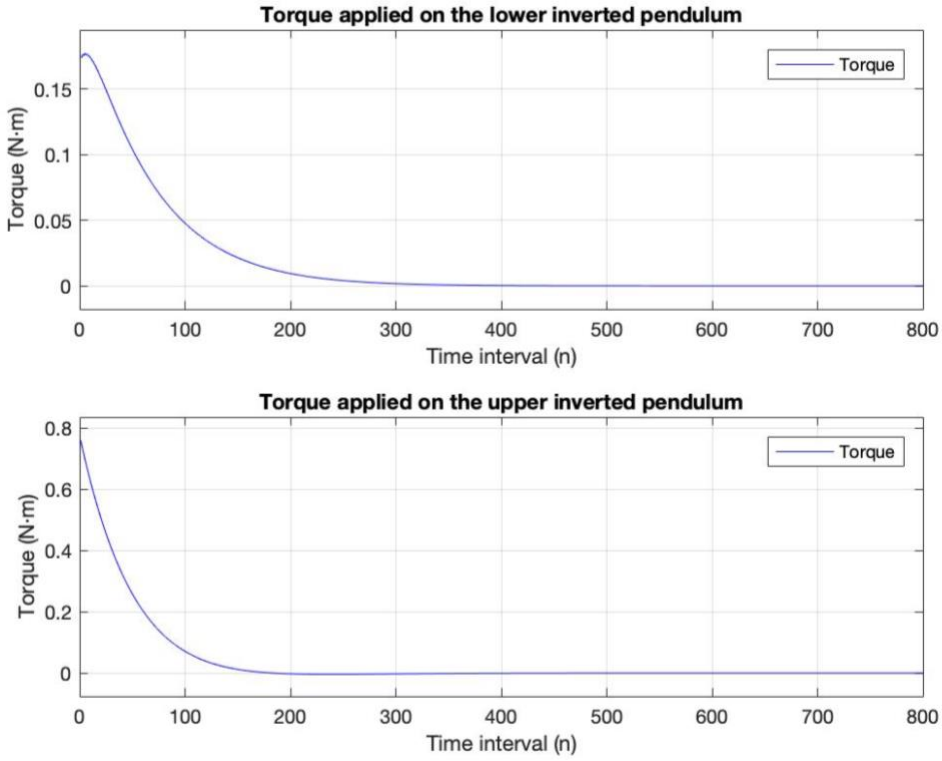

Figure 5. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1773 N·m.

The maximum torque of the upper motor is 0.7603 N·m and the minimum torque of the upper motor is -0.0043 N·m.

4.1.2 Feedback Gain K_1 = [64 1.5 23 3.2]

1 Co-learning model

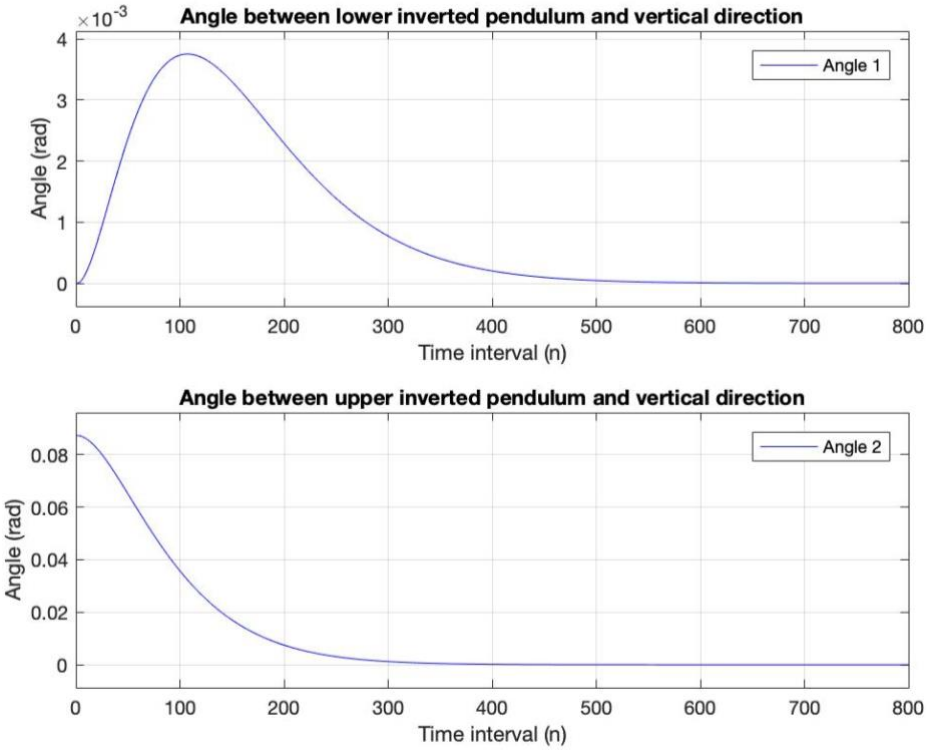

Figure 6. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0038 Rad at time n = 107 time intervals.

It takes 495 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 461 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

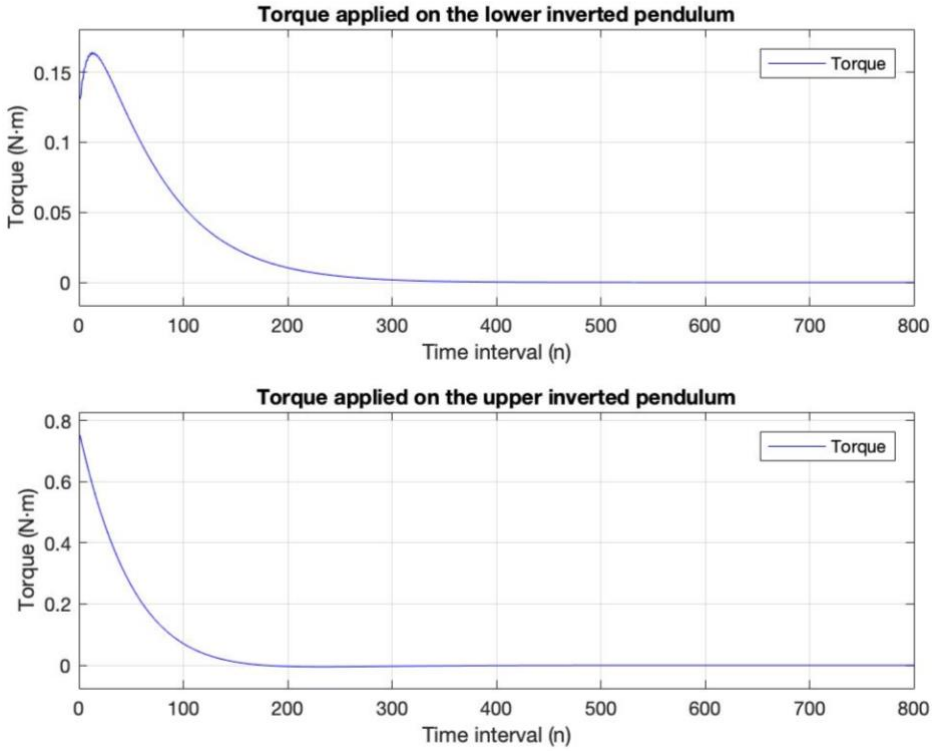

Figure 7. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1642N·m.

The maximum torque of the upper motor is 0.7520 N·m and the minimum torque of the upper motor is -0.0051 N·m.

2. No learning ability model

The initial control strategy of human K_1=[64 1.5 23 3.2]. The corresponding machine control strategy can be obtained as K_2=[2.87 8.62 2.27 2.60].

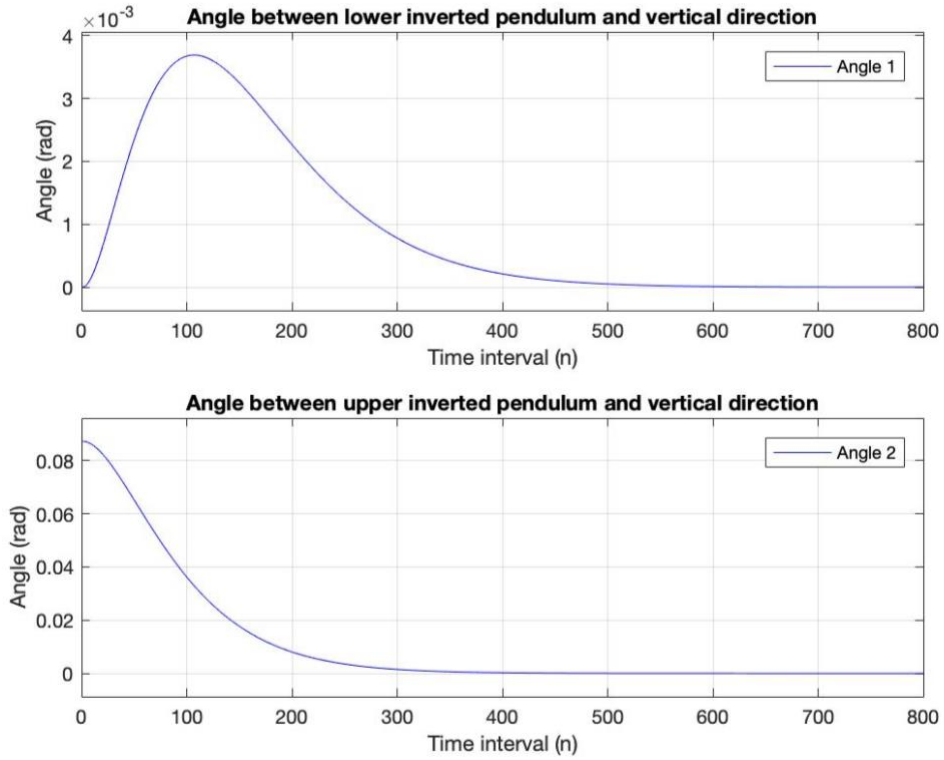

Figure 8. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0037 Rad at time n = 107 time intervals.

It takes 501 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 504 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

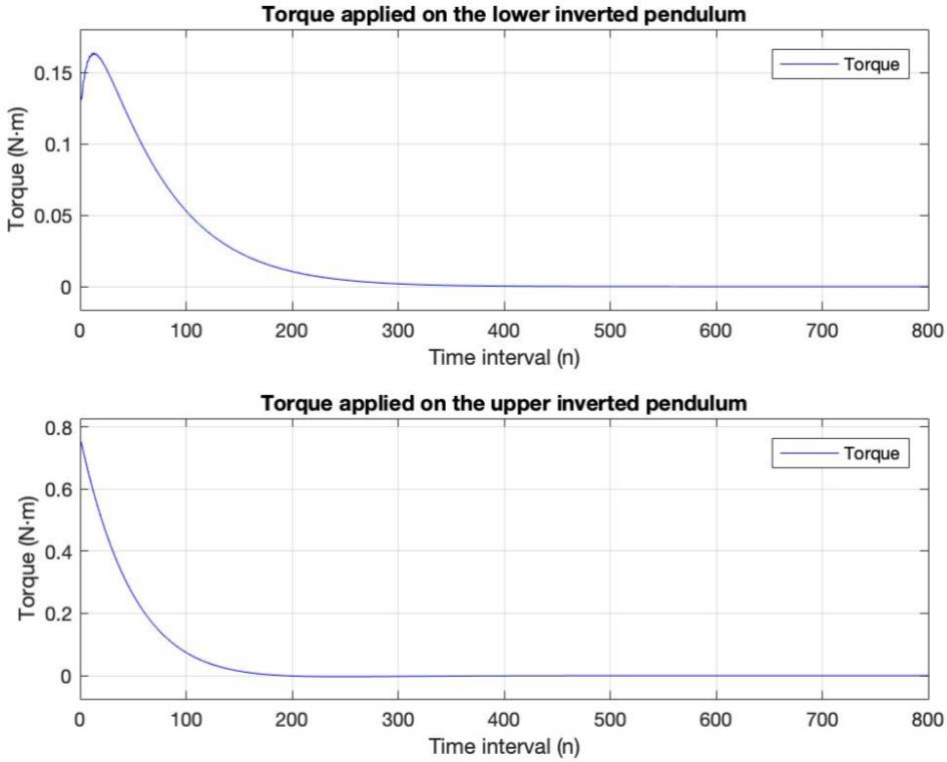

Figure 9. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1640N·m.

The maximum torque of the upper motor is 0.7520 N·m and the minimum torque of the upper motor is -0.0034 N·m.

4.1.3 Feedback Gain K_1 = [126 0.5 64 8.2]

1 Co-learning model

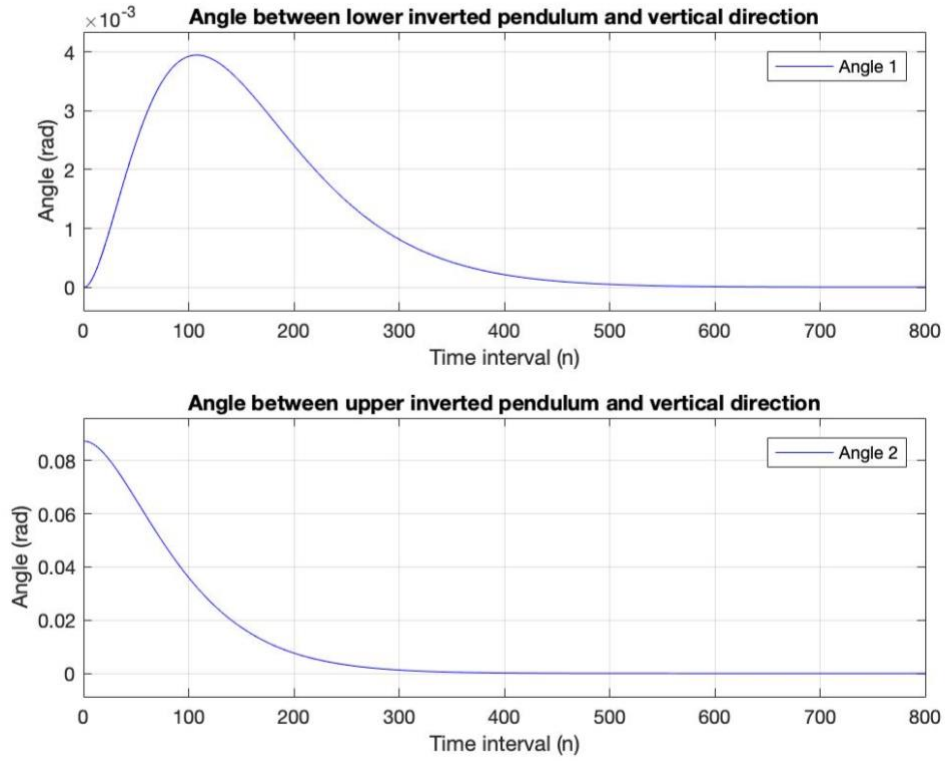

Figure 10. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0040 Rad at time n = 108 time intervals.

It takes 499 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 465 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

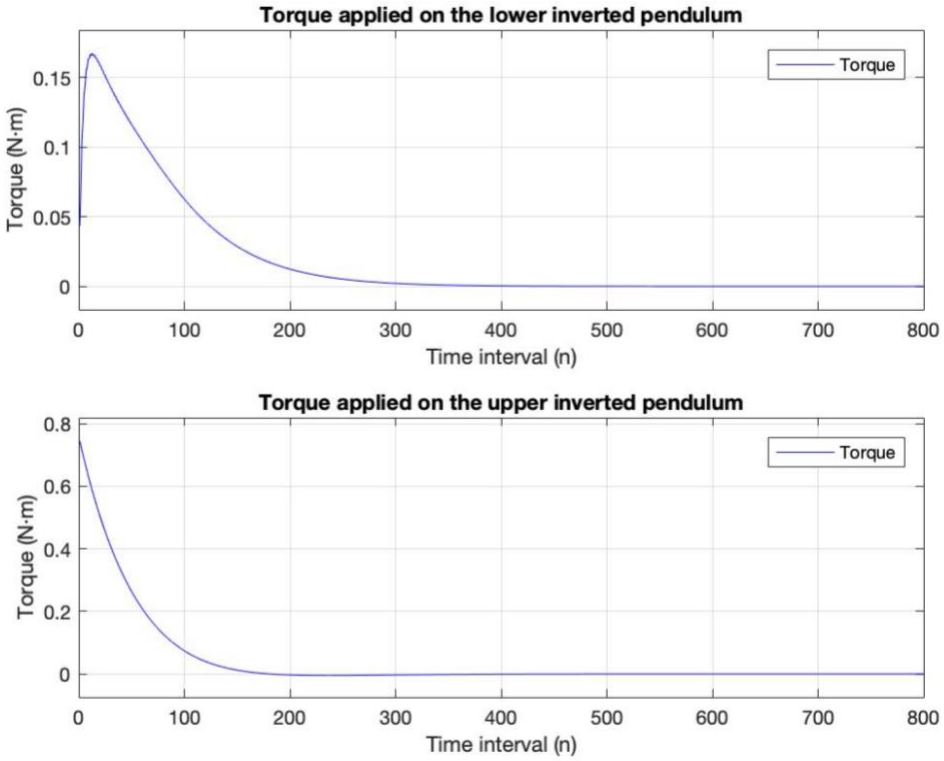

Figure 11. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1675N·m.

The maximum torque of the upper motor is 0.7446 N·m and the minimum torque of the upper motor is -0.0050 N·m.

2. No learning ability model

The initial control strategy of human K_1=[126 0.5 64 8.2]. The corresponding machine control strategy can be obtained as K_2=[2.84 8.54 2.49 2.62].

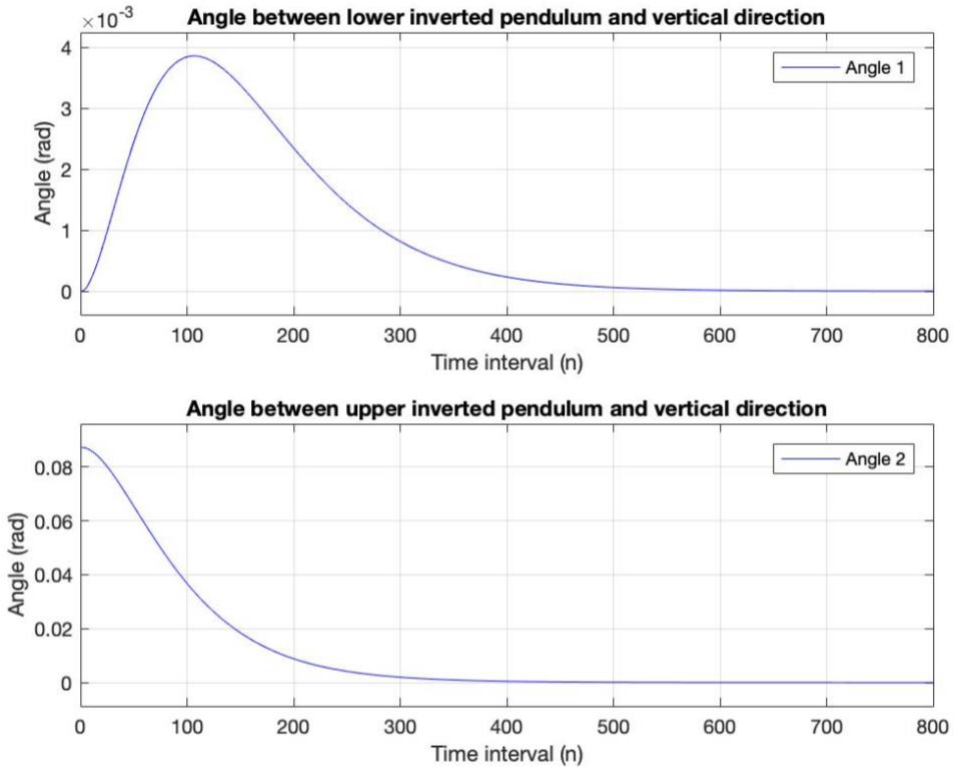

Figure 12. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0039 Rad at n = 107 time intervals.

It takes 515 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 569 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

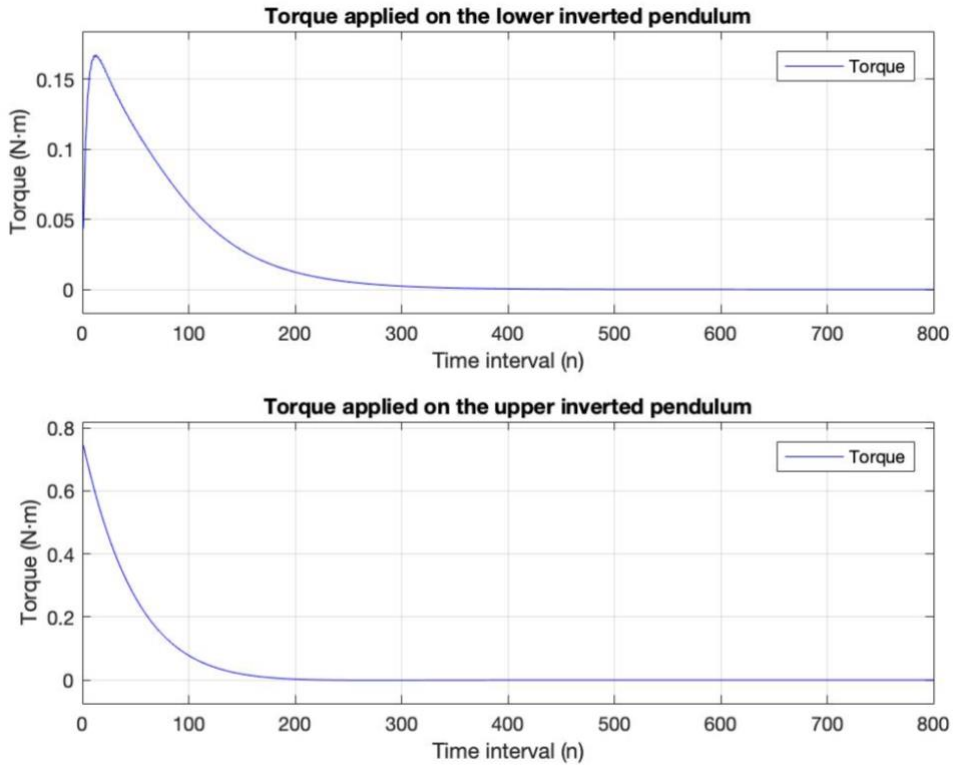

Figure 13. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1673N·m.

The maximum torque of the upper motor is 0.7446 N·m and the minimum torque of the upper motor is -9.4787· 10^{-4} N·m.

4.2 Changing the Weighting Matrix *Q* and *R*

Assuming the initial feedback gain *K* and the update rate α are constants. Comparing the simulation results using Co-learning model when the weighting matrix *Q* and *R* changed.

In the simulation below, let K_1=[45 2 15 2.2], updating rate α = 0.1.

4.2.1 Weighting Matrix *Q* = *diag* [1 , 1 , 0 , 0], *R* = 1

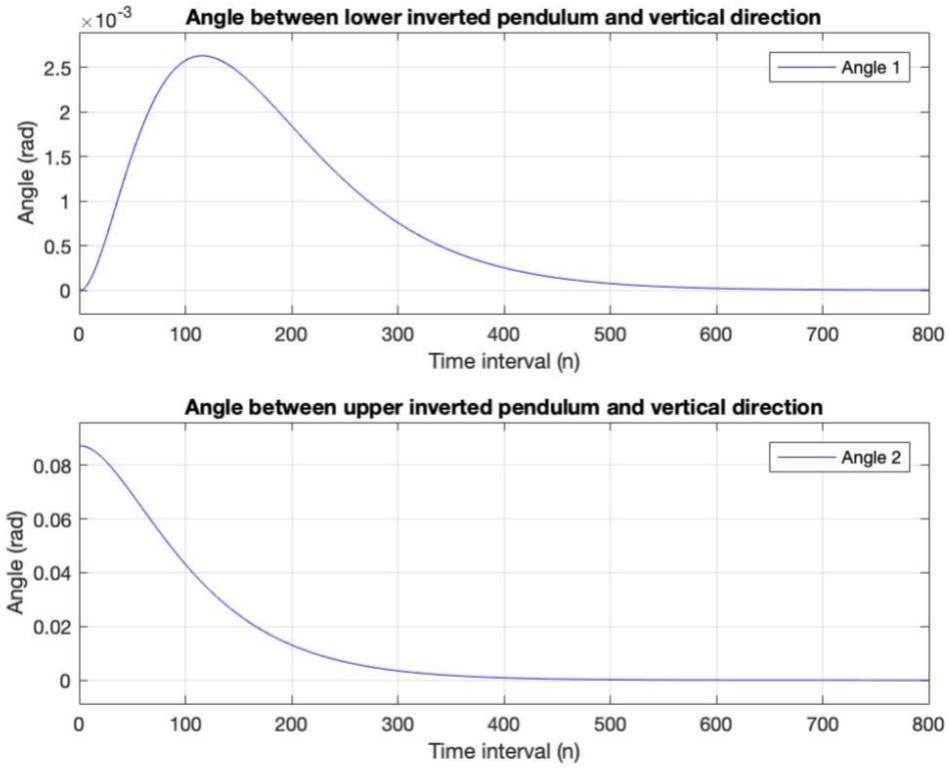

Figure 14. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0026 Rad at time n = 116 time intervals.

It takes 532 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 607 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

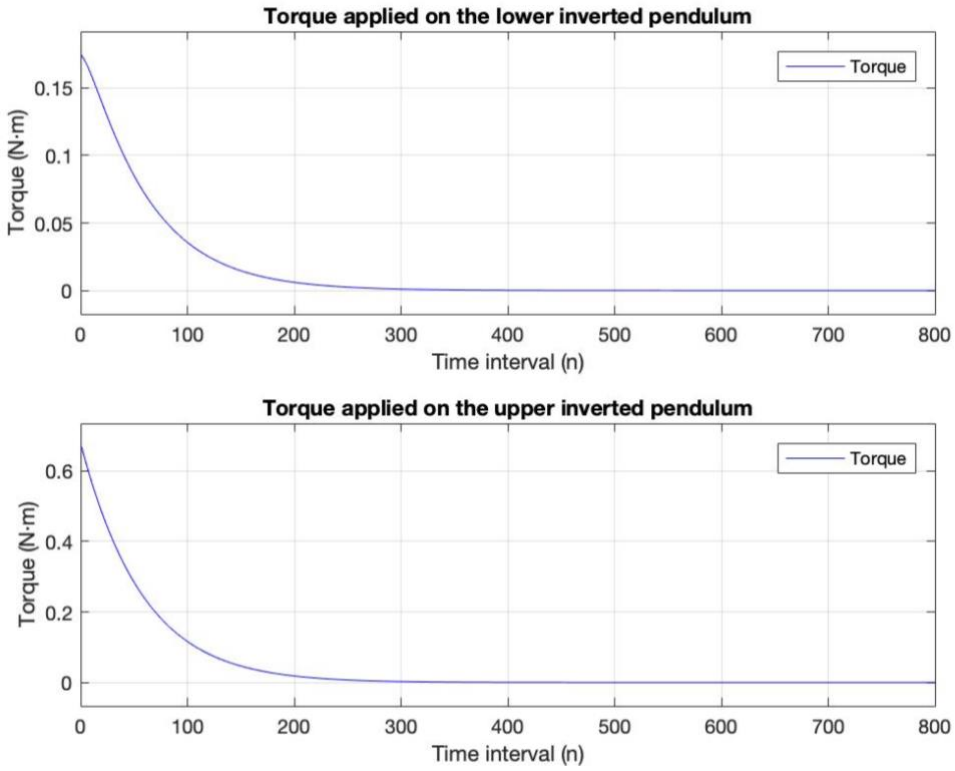

Figure 15. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1744N·m.

The maximum torque of the upper motor is 0.6685 N·m and the minimum torque of the upper motor is -1.2090· 10^{-5} N·m.

4.2.2 Weighting Matrix $Q = diag\ [10\ ,\ 10\ ,\ 0\ ,\ 0],\ R = 0.1$

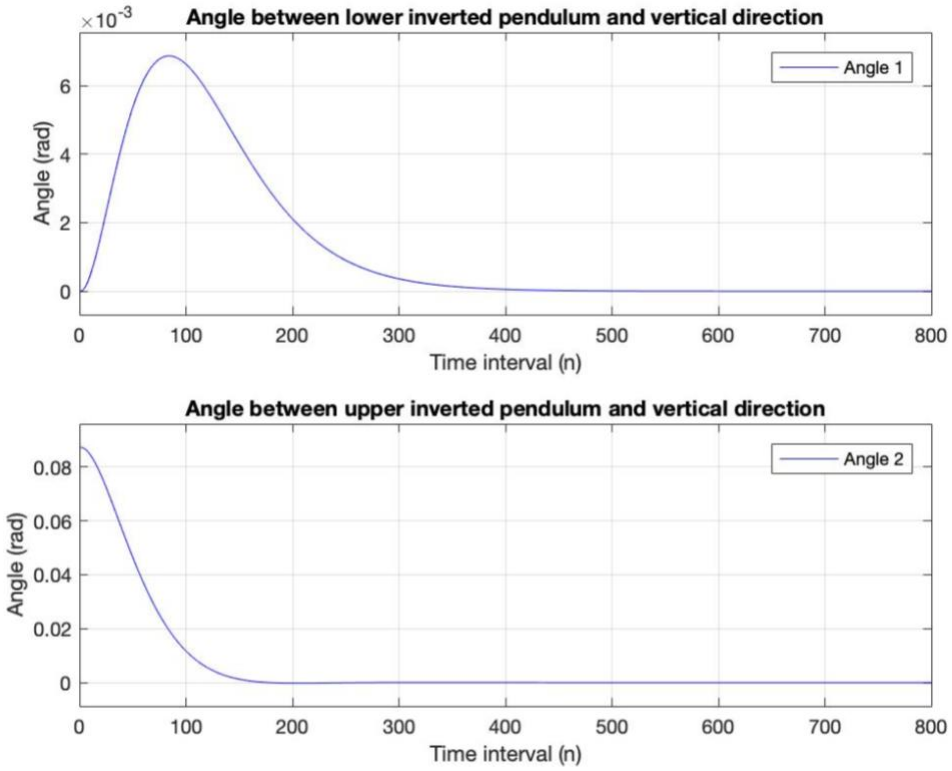

Figure 16. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0069 Rad at time n = 84 time intervals.

It takes 404 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 257 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

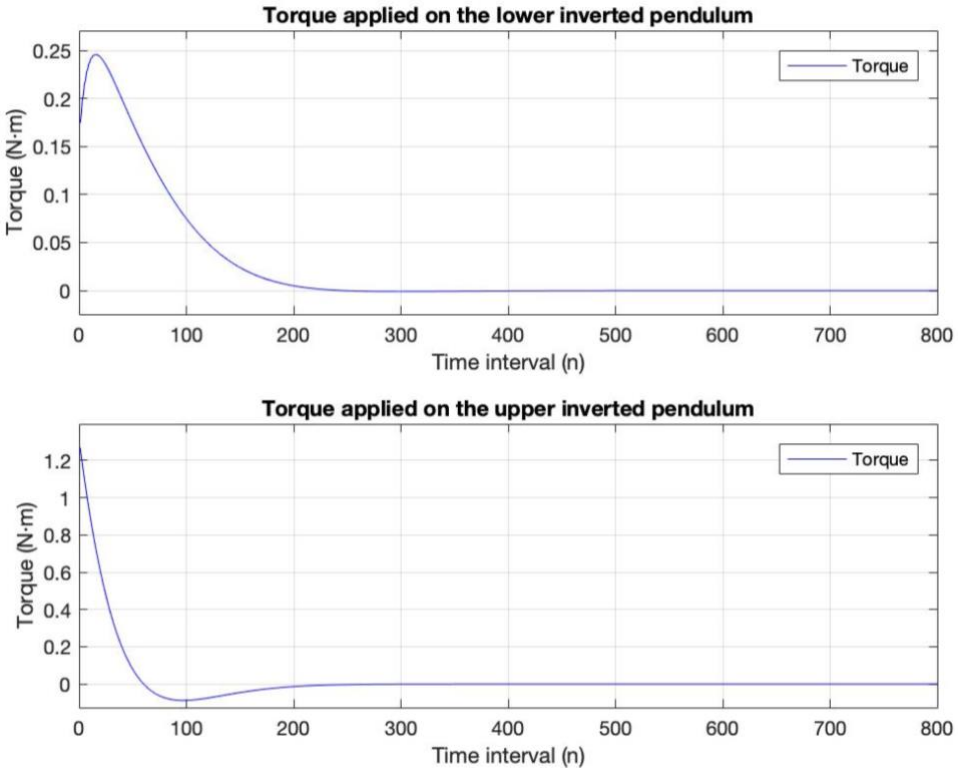

Figure 17. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.2461 N·m.

The maximum torque of the upper motor is 1.2692 N·m and the minimum torque of the upper motor is -0.0885 N·m.

4.3 Changing the Updating Rate α

Assuming the state weighting matrix Q, input weighting matrix R and the initial feedback gain K are constants. Comparing the simulation results when the updating rate α changed.

In the simulation below, let K_1=[126 0.5 64 8.2], Q=*diag*[10 , 10 , 0 , 0], R=1.

4.3.1 Updating Rate α = 0.5

1 Co-learning model

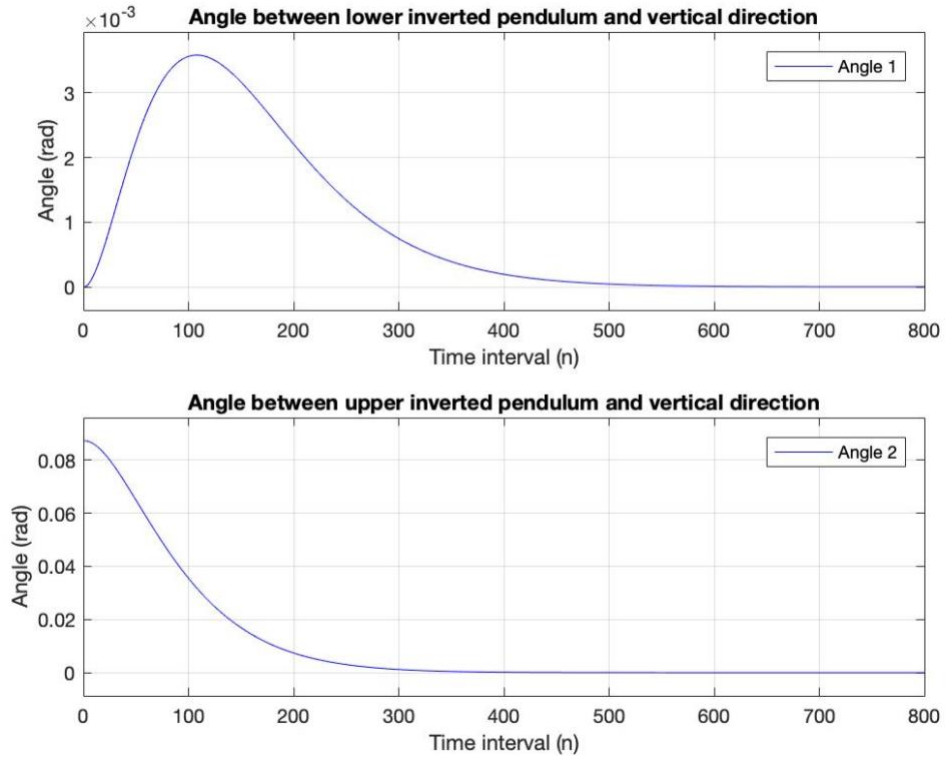

Figure 18. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0036 Rad at time n = 108 time intervals.

It takes 492 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 458 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

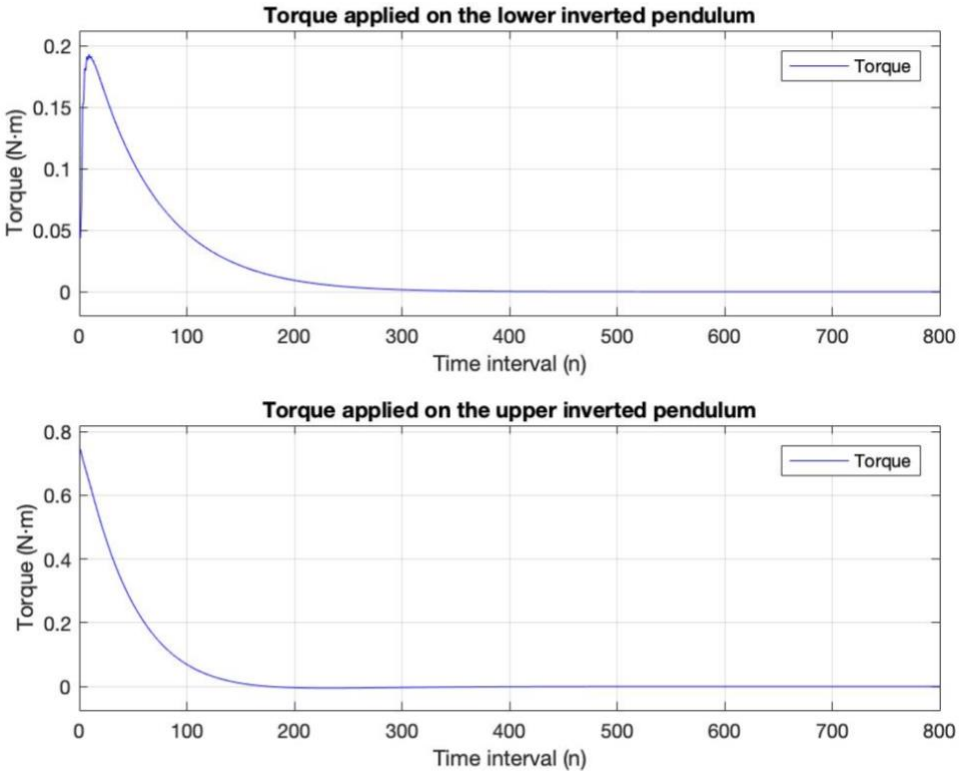

Figure 19. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1929 N·m.

The maximum torque of the upper motor is 0.7446 N·m and the minimum torque of the upper motor is -0.0052 N·m.

2 No learning ability model

The initial control strategy of human K_1=[126 0.5 64 8.2]. The corresponding machine control strategy can be obtained as K_2=[2.84 8.54 2.49 2.62].

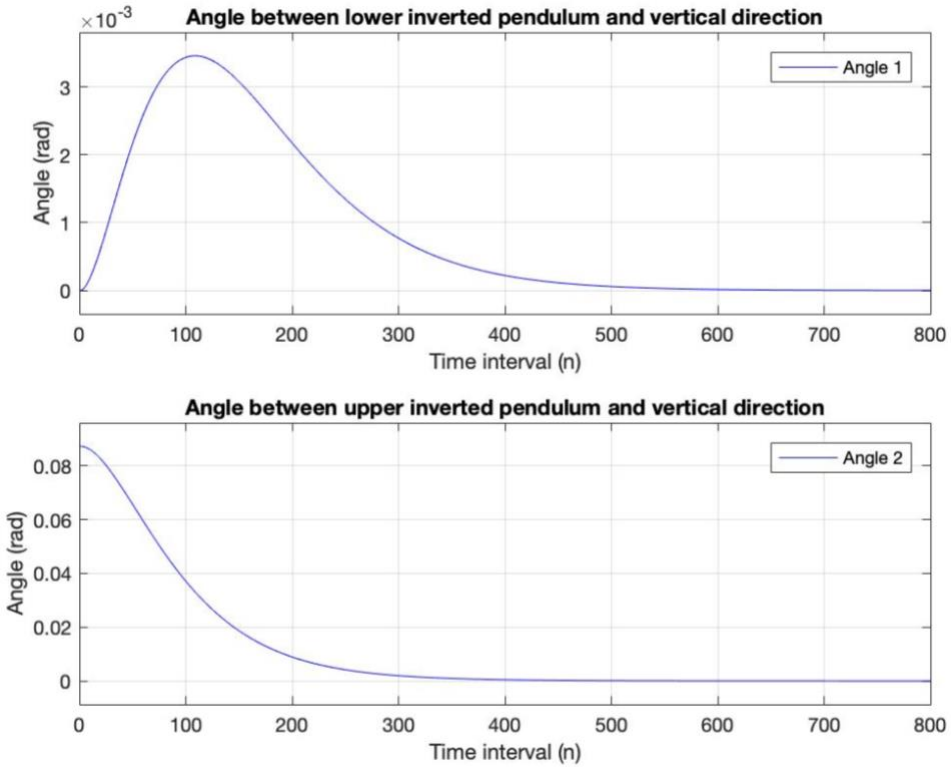

Figure 20. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0035 Rad at n = 109 time intervals.

It takes 511 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 565 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

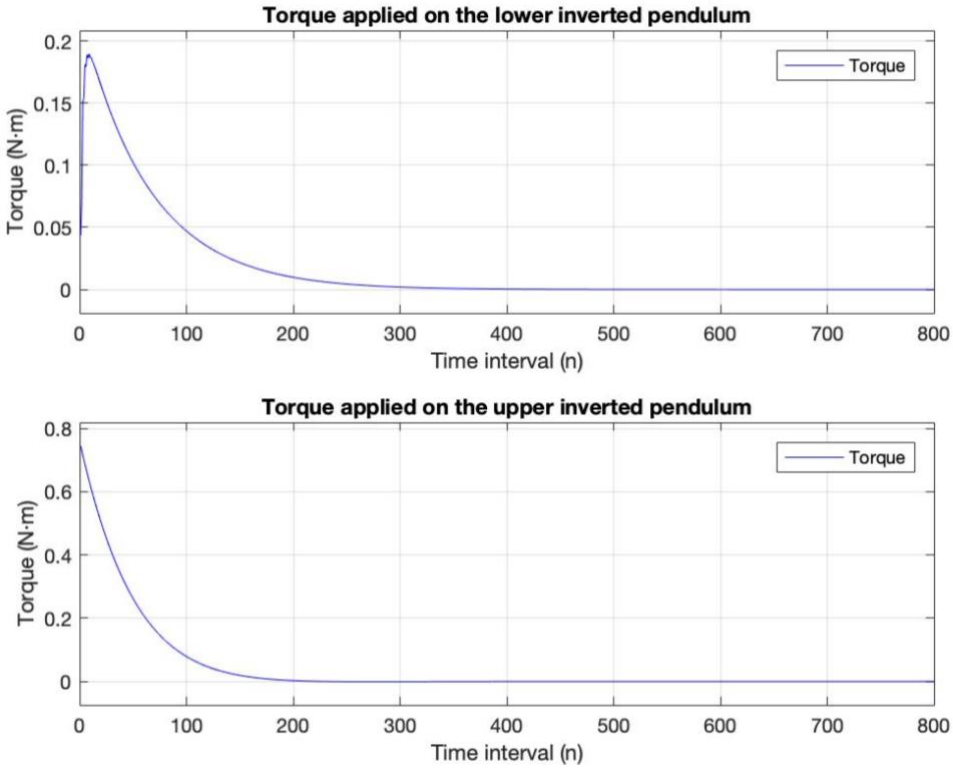

Figure 21. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1894 N·m.

The maximum torque of the upper motor is 0.7446 N·m and the minimum torque of the upper motor is -0.0010 N·m.

4.4 Time Delay

It is assumed that the machine can apply the updated control strategies to the system in real time. For humans, there will be a delay from obtaining the new control strategies to applying the strategies to the system.

Assuming the state weighting matrix Q, input weighting matrix R, the initial feedback gain K and updating rate α are constants. Comparing the simulation results using Co-learning model when the time delay τ changed.

In the simulation below, let K_1=[126 0.5 64 8.2], Q=*diag*[10 , 10 , 0 , 0], R=1, $\alpha = 0.1$.

4.4.1 Time Delay τ = 10 Time Intervals

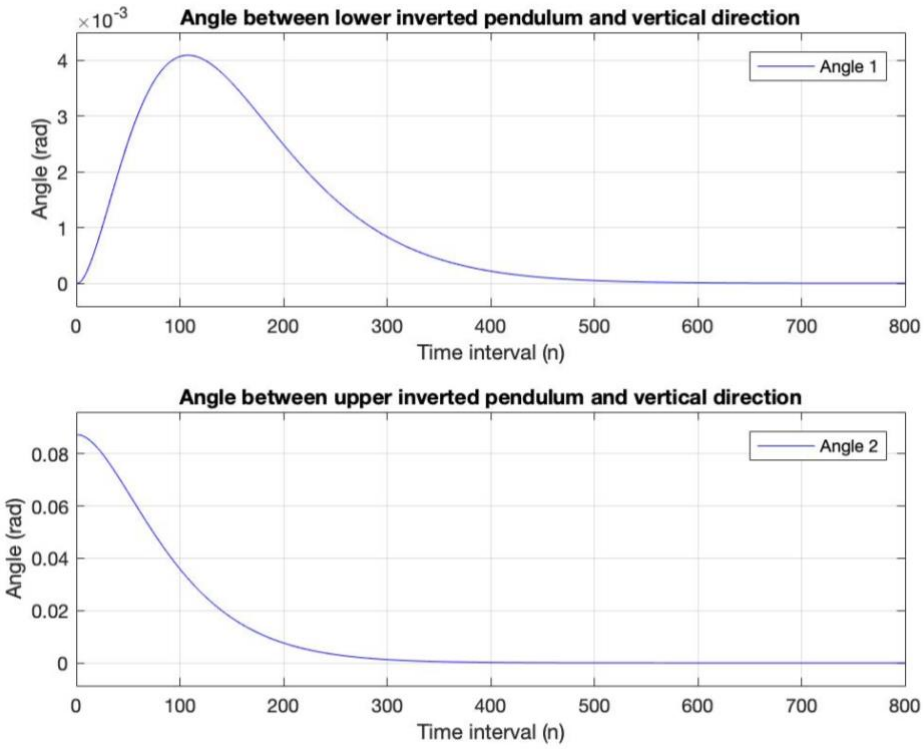

Figure 22. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0041 Rad at time n = 108 time intervals.

It takes 500 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 467 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

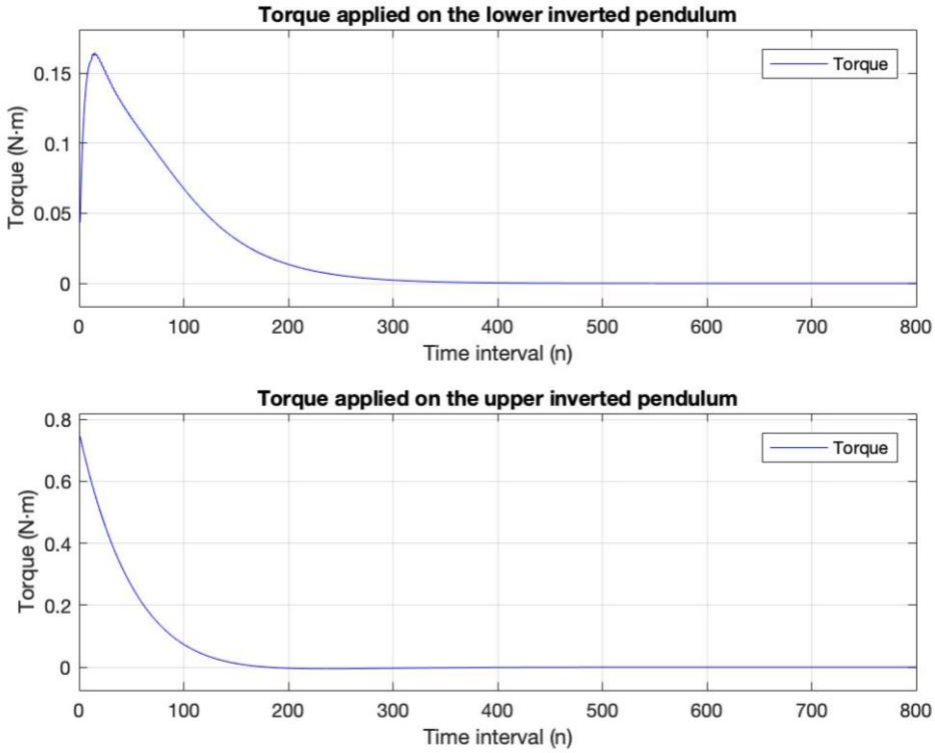

Figure 23. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1647 N·m.

The maximum torque of the upper motor is 0.7446 N·m and the minimum torque of the upper motor is -0.0050 N·m.

4.4.2 Time Delay τ = 100 Time Intervals

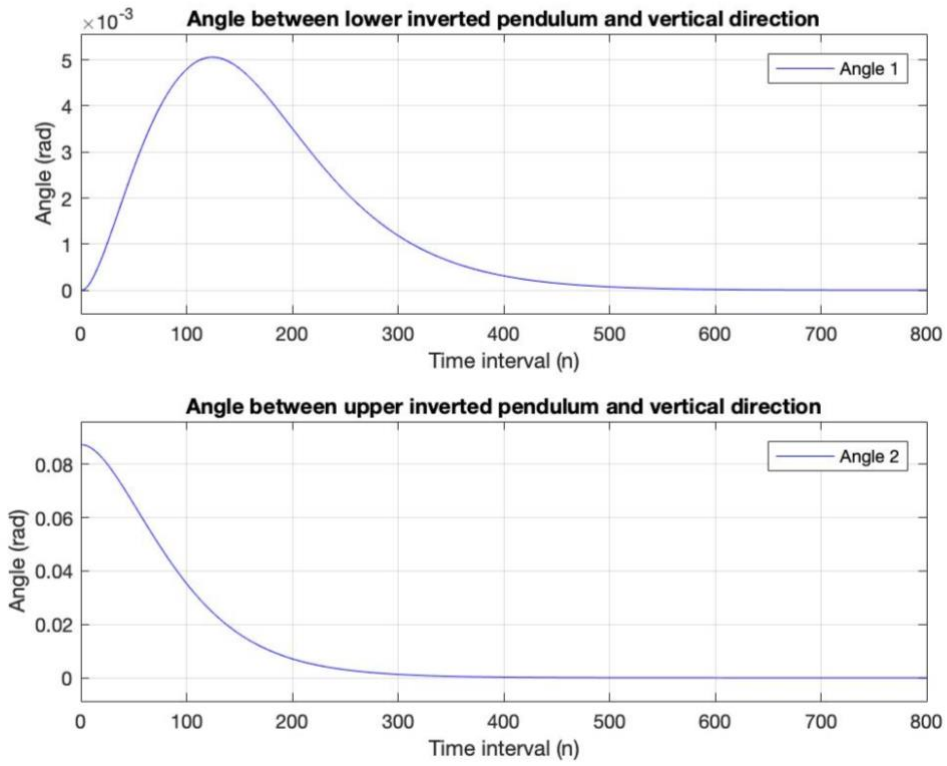

Figure 24. Angle between inverted pendulums and vertical direction.

The lower inverted pendulum reaches the maximum angle 0.0051 Rad at time n = 124 time intervals.

It takes 525 time intervals for $\angle 1$ to converge below $5 \cdot 10^{-5}$ radians and it takes 491 time intervals for $\angle 2$ to converge below $5 \cdot 10^{-5}$ radians.

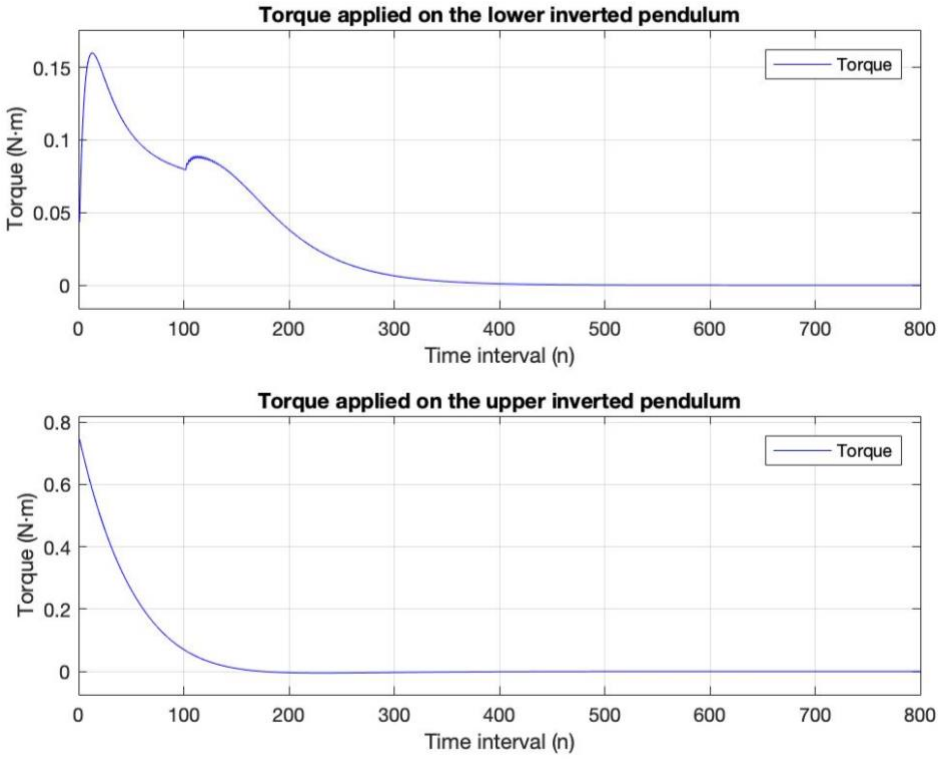

Figure 25. Torque applied on the pendulums.

The maximum torque of the lower motor is 0.1598 N·m.

The maximum torque of the upper motor is 0.7446 N·m and the minimum torque of the upper motor is -0.0051 N·m.

4.5 Simulation Results

The above simulation results can be summarized into the following tables.

Table 2. System performance when changing the initial feedback gain *K*

	K_1		K_2		K_3	
	Co-learning	No learning ability	Co-learning	No learning ability	Co-learning	No learning ability
N_1	492	496	495	501	499	515
N_2	458	483	461	504	465	569
$\angle_1$ Max	0.0036	0.0036	0.0038	0.0037	0.0040	0.0039
u_1Max	0.1773	0.1773	0.1642	0.1640	0.1675	0.1673
u_2Max	0.7603	0.7603	0.7520	0.7520	0.7446	0.7446
u_2Min	-0.0052	-0.0043	-0.0051	-0.0034	-0.0050	-0.0009

From results above, it can be observed that

1. For the same model, the greater the difference between the initial feedback gain and the calculated optimal solution, the longer the time to converge to a fixed angle range.
2. Different models have significant differences in convergence time even if they have the same initial feedback gain. Since both human and machine have the ability to learn, there is no significant difference in convergence time of the co-learning model. In the second

model, the machine does not have the learning ability, and the control strategy cannot be optimized, so the time to converge to a fixed angle range is significantly increased.

3. The difference between the inverted pendulum and the vertical direction represents the state error, while the torque represents the energy entering the system. Since the weight matrices *Q* and *R* are fixed, the state error and the energy entering the system do not change significantly, and the maximum and minimum values remain unchanged.

Table 3. System performance when changing the weighting matrix *Q* and *R*

	$Q_1\ R_1$	$Q_2\ R_1$	$Q_1\ R_2$
N_1	492	532	404
N_2	458	607	257
$\angle_1$ Max	0.0036	0.0026	0.0069
u_1Max	0.1773	0.1744	0.2461
u_2Max	0.7603	0.6685	1.2692
u_2Min	-0.0052	-1.2090e-05	-0.0885

From results above, it can be observed that

1. Under the condition that the initial feedback coefficient *K*, the update rate and the *R* matrix remain unchanged, when the *Q* matrix is reduced, the maximum angle of the lower inverted pendulum is reduced, which means that the state error of the system is

reduced and the system is more stable. But the time for the system to reach steady state has increased.

2. Under the condition that the feedback parameter *K*, the update rate and the *Q* matrix are unchanged, as the *R* matrix decreases, the maximum torque generated by the motor increases, meaning that the energy input into the system increases. At the same time, the maximum angle of the inverted pendulum increases and the system becomes more difficult to stabilize. However, the convergence time becomes faster.

   The simulation results are completely consistent with the properties of the LQR controller introduced earlier.

Table 4. System performance when changing the updating rate α

	α_1		α_2	
	Co-learning	No learning ability	Co-learning	No learning ability
N_1	499	515	492	511
N_2	465	569	458	565
$\angle_1$ Max	0.0040	0.0039	0.0036	0.0035
u_1Max	0.1675	0.1673	0.1929	0.1894
u_2Max	0.7446	0.7446	0.7446	0.7446
u_2Min	-0.0050	-0.0009	-0.0052	-0.0010

From results above, it can be observed that when the update rate increases, the convergence time of both models decreases, but the convergence rate of the co-learning model is faster than the convergence rate of the non-learning model. Since the machine does not have the learning ability and the control strategy cannot be updated in the second model, the control strategy applied to the upper inverted pendulum is not the optimal control strategy. Even if the update speed is increased, it is updated to the non-optimal control strategy as soon as possible. Therefore, the convergence rate is slower than the co-learning model.

Table 5. System performance with time delay

Time delay	0	10	100
N_1	499	500	525
N_2	465	467	491
$\angle_1$ Max	0.0040	0.0041	0.0051
u_1Max	0.1675	0.1647	0.1598
u_2Max	0.7446	0.7446	0.7446
u_2Min	-0.0050	-0.0050	-0.0051

From results above, it can be observed that as the delay time for humans to apply the new control strategy to the system increases, the time for the upper and lower inverted pendulum to converge to a certain fixed range increases, and the stability of the system deteriorates. Similarly,

since the Q matrix and the R matrix are unchanged, the maximum angle of the inverted pendulum and the maximum torque of the motor do not change significantly.

5.0 Conclusion and Future Work

5.1 Conclusion

In this thesis, a second-order inverted pendulum system is designed which is controlled by human and machine corporately. A co-learning controller using LQR technique is introduced. The simulation results indicate that the co-learning controller converge faster than controller that do not have the learning ability. At the same time, the initial control feedback gain, iterative updating rate and the weighting matrix Q and R are adjusted in the simulation to get the performances of the co-learning controller.

5.2 Future Work

In this thesis, only simulations have been performed for the co-learning controllers. Future studies will focus on the experiment results.

Besides, in this paper, it is assumed that the people involved in the experiment are trained and humans are abstracted into a LQR controller, so the human control strategies are optimal. A more specific optimal model of human response was introduced in[5], [6]. The model of human response can be split into time delay τ, equalization network and neuromotor dynamics with observation noise V_y(t) and motor noise V_u (t). Therefore, the model of manual control needs to be further improved.

Another direction for future work is to improve the controller. In this paper, the different parameters of the LQR controller are adjusted, which qualitatively proves the feasibility of co-learning controller and the influence of different parameters on control performance. Machine learning, neural network and other technologies which will make the controller more close to human response will be added to the controller.

Appendix A Matlab Code

This Matlab code is applicable to the case of K_1=[126 0.5 64 8.2], human-machine co-learning controller

```
m1=1.5;
m2=0.5;
m3=0.75;
L1=1;
L2=1.5;
l1=0.5;
l2=0.75;
g=9.8
a1=4/3*m1*l1^2+4*m2*l1^2+4*m3*l1^2;
a2=2*m2*l1*l2;
a3=2*m2*l1*l2;
a4=4/3*m2*l2^2;
f1=m1*g*l1+2*m3*g*l1+2*m2*g*l1;
f2=m2*g*l2;
a=[1 0 0 0; 0 1 0 0; 0 0 a1 a2; 0 0 a3 a4];
b=[0 0 1 0; 0 0 0 1; f1 0 0 0; 0 f2 0 0];
c=[0 0; 0 0; 1 0; 0 1];
A=inv(a)*b;
B=inv(a)*c;
C=[1 0 0 0; 0 1 0 0];
D=[0 0; 0 0];
B1=B(:,1);
B2=B(:,2);
K1=[126 0.5 64 8.2];
A2=[(A-B1*K1)];
B2=B2;
Q=[10 0 0 0; 0 10 0 0; 0 0 0 0; 0 0 0 0];
R=1;
[K2,P2]=lqr(A2,B2,Q,R);
N = 2000
x = zeros(4,N)
u=zeros(2,N)
T=0.005
x(2, 1)=5/180*3.14
for i=1:N
K=[K1;K2];
u(1,i)=K1*x(:,i);
u(2,i)=K2*x(:,i);
Ac=[(A-B*K)];
Bc=[B];
Cc=[C];
Dc=[D];
```

```
[Ad, Bd] = c2d(Ac, Bc, T)
x(:,i+1)=Ad*x(:,i);
if mod(i,2)==0;
A1=[(A-B2*K2)];
B1=B1;
Q1=[10 0 0 0; 0 10 0 0; 0 0 0 0; 0 0 0 0];
R1=1;
[K1s,P1]=lqr(A1,B1,Q1,R1);
alpha1=0.5;
K1=K1+alpha1*(K1s-K1);
end
A2=[(A-B1*K1)];
B2=B2;
Q2=[10 0 0 0; 0 10 0 0; 0 0 0 0; 0 0 0 0];
R2=1;
[K2s,P2]=lqr(A2,B2,Q2,R2);
alpha2=0.5;
K2=K2+alpha2*(K2s-K2);
end
figure(1)
subplot(2,1,1);
plot(x(1,:), 'b');
title('Angle between lower inverted pendulum and vertical direction');
axis([0 800 -0.1*max(x(1,:)) 1.1*max(x(1,:))])
ylabel('Angle (rad)')
xlabel('Time interval (n)')
grid;
[ymax1,tmax1]=max(x(1,:))
[ymin1,tmin1]=min(x(1,:))
legend('Angle 1')
subplot(2,1,2);
plot(x(2,:), 'b');
title('Angle between upper inverted pendulum and vertical direction');
axis([0 800 -0.1*max(x(2,:)) 1.1*max(x(2,:))])
ylabel('Angle (rad)')
xlabel('Time interval (n)')
grid;
[ymax2,tmax2]=max(x(2,:))
[ymin2,tmin2]=min(x(2,:))
legend('Angle 2')
figure(2)
subplot(2,1,1);
plot(u(1,:),'b-')
title('Torque applied on the lower inverted pendulum');
axis([0 800 -0.1*max(u(1,:)) 1.1*max(u(1,:))])
ylabel('Torque (N∑m)')
xlabel('Time interval (n)')
grid;
[umax1,tmax1]=max(u(1,:))
[umin1,tmin1]=min(u(1,:))
legend('Torque');
subplot(2,1,2);
plot(u(2,:),'b-')
title('Torque applied on the upper inverted pendulum');
axis([0 800 -0.1*max(u(2,:)) 1.1*max(u(2,:))])
ylabel('Torque (N∑m)')
```

```
xlabel('Time interval (n)')
grid;
[umax2,tmax2]=max(u(2,:))
[umin2,tmin2]=min(u(2,:))
legend('Torque');
```

Bibliography

[1] T. Shmelova, Y. Sikirda, N. Rizun, V. Lazorenko, and V. Kharchenko, "Machine Learning and Text Analysis in an Artificial Intelligent System for the Training of Air Traffic Controllers," 2019, pp. 1–50.

[2] J. Hua, Y. Cui, P. Shi, Y. Yang, and H. Li, "Force feedback assisted balancing of inverted pendulum under manual control," *Proc. 2013 Int. Conf. Intell. Control Inf. Process. ICICIP 2013*, pp. 776–781, 2013.

[3] M. Lupu, M. Sun, D. Askey, R. Xia, and Z. H. Mao, "Human strategies in balancing an inverted pendulum with time delay," *2010 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC'10*, no. 1, pp. 5246–5249, 2010.

[4] N. Amirshirzad, A. Kumru, and E. Oztop, "Human adaptation to human-robot shared control," *IEEE Trans. Human-Machine Syst.*, vol. 49, no. 2, pp. 126–136, 2019.

[5] D. L. Kleinman, S. Baron, and W. H. Levison, "An optimal control model of human response part I: Theory and validation," *Automatica*, vol. 6, no. 3, pp. 357–369, 1970.

[6] S. Baron, D. L. Kleinman, and W. H. Levison, "An optimal control model of human response part II: prediction of human performance in a complex task," *Automatica*, vol. 6, no. 3, pp. 371–383, 1970.

[7] D. P. Lindorff, "Relay control of systems with parameter uncertainties and disturbances," *Automatica*, vol. 5, no. 6, pp. 755–762, 2003.

[8] W. R. Sturgeon and W. V Loscutoff, "Application of modal control and dynamic observers to control of a double inverted pendulum," *Jt. Autom. Control Conf.*, vol. 10, pp. 857–865, 1972.

[9] E. E. Eldukhri and M. Cetin, "Balancing and attitude control," no. 3, 1995.

[10] H. Li and D. Ceglarek, "Optimal Trajectory Planning For Material Handling of Compliant Sheet Metal Parts," *J. Mech. Des.*, vol. 124, no. 2, p. 213, 2002.

[11] R. Orostica, M. A. Duarte-Mermoud, and C. Jauregui, "Stabilization of inverted pendulum using LQR, PID and fractional order PID controllers: A simulated study," pp. 1–7, 2016.

[12] E. S. Varghese, A. K. Vincent, and V. Bagyaveereswaran, "Optimal control of inverted pendulum system using PID controller, LQR and MPC," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 263, no. 5, 2017.

[13] C. H. Chiu, “The design and implementation of a wheeled inverted pendulum using an adaptive output recurrent cerebellar model articulation controller,” *IEEE Trans. Ind. Electron.*, vol. 57, no. 5, pp. 1814–1822, 2010.

[14] D. Q. Mayne and J. B. Rawlings, “Correction to ‘Constrained model predictive control: stability and optimality,’” *Automatica*, vol. 37, no. 3, p. 483, 2002.

[15] M. Askari, H. A. F. Mohamed, M. Moghavvemi, and S. S. Yang, “Model predictive control of an inverted pendulum,” *2009 Int. Conf. Tech. Postgraduates*, pp. 1–4, 2010.

[16] J. T. Wang, G. P. Liu, W. Jin, and G. F. Xiao, “The Application of Double Closed-Loop Unit Gain Negative Feedback and BP Neural Network Controller in Single Inverted Pendulum,” *Adv. Mater. Res.*, vol. 490–495, pp. 1723–1727, Mar. 2012.

[17] S. Jung and S. S. Kim, “Control Experiment of a Wheel-Driven Mobile Inverted Pendulum Using Neural Network,” *IEEE Trans. Control Syst. Technol.*, vol. 16, no. 2, pp. 297–303, 2008.

[18] V. Mohan, A. Rani, and V. Singh, “Robust adaptive fuzzy controller applied to double inverted pendulum,” *J. Intell. Fuzzy Syst.*, vol. 32, no. 5, pp. 3669–3687, 2017.

[19] C. Luo, D. Hu, Y. Pang, X. Zhu, and G. Dong, “Fuzzy control of a quintuple inverted pendulum with the LQR method and 2-ary fuzzy piecewise interpolation function,” pp. 6307–6312, 2008.

[20] H. M. J. and N. B. N. and R. S. Phillips, *Theory of Servomechanisms. (1st ed.)*. New York: McGraw-Hill Book Co, 1947.

[21] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*. Cambridge: Technology Press of the Massachusetts Institute of Technology, 1949.

[22] A. Tustin, *Journal of the Institution of Electrical Engineers - Part IIA: Automatic Regulators and Servo Mechanisms*. The Institution of Engineering and Technology, 1947.

[23] U. A. D. Number *et al.*, “Unclassified ad number,” *Test*, no. May, 1991.

[24] J. I. Elkind, “Characteristics of simple manual control systems.” Massachusetts Institute of Technology, 1956.

[25] D. T. McRuer and E. S. Krendel, “Dynamic response of human operators,” KELSEY-HAYES CO INGLEWOOD CA CONTROL SPECIALISTS DIV, 1957.

[26] D. T. McRuer, W. Reisner, E. S. Krendel, and D. Graham, *Human Pilot Dynamics in Compensatory Systems; Theory, Models, and Experiments with Controlled Element and Forcing Function Variation*. 1965.

[27] X. Chen, R. Yu, K. Huang, S. Zhen, H. Sun, and K. Shao, “Linear motor driven double inverted pendulum: A novel mechanical design as a testbed for control algorithms,” *Simul. Model. Pract. Theory*, vol. 81, pp. 31–50, 2018.

[28] Y. Q. Jin, J. W. Lei, and D. Liu, “Modeling and PID Control of Single-Stage Inverted Pendulum System,” *Appl. Mech. Mater.*, vol. 644–650, pp. 142–145, Sep. 2014.

[29] M. Venturin and S. Poles, “Controller for an Inverted Pendulum,” p. 32.

[30] Y. Sun, J. Lyu, J. Fang, Z. Fu, and D. Dong, “Robust LQR Anti-Swing Control for Quay-Side Crane System with Variable Load,” *2018 IEEE 8th Annu. Int. Conf. CYBER Technol. Autom. Control. Intell. Syst.*, pp. 796–801, 2019.

[31] K. Miyamoto, J. She, D. Sato, and N. Yasuo, “Automatic determination of LQR weighting matrices for active structural control,” *Eng. Struct.*, vol. 174, no. April, pp. 308–321, 2018.