'89 KACC 1989. 10. 27~28

Composite Neural Networks for Controlling Semi-Linear Dynamical Systems: Example from Inverted Pendulum Problem

Yoshinobu Yamamoto Yuichiro Anzai
Department of Computer Science
Graduate School of Science and Technology
Keio University
3-14-1, Hiyoshi
Yokohama 223 Japan

1 Introduction

In this paper, we propose a neural network for learning to control semi-linear dynamical systems. The network is a composite system of four three-layer backpropagation subnetworks, and is able to control inverted pendulums better than systems based on modern control theory at least in some ranges of parameters.

Three of the four subnetworks in our network system process angles, velocities, and positions of a moving inverted pendulum, respectively. The outputs from those three subnetworks are input to the remaining subnetwork that makes control decisions. Each of the four subnetworks learns connection weights independently by backpropagation algorithms. Teaching signals are given by the human operator. Also, input signals are generated by the human operator, but they are converted by preprocessors to actual input data for the three subnetworks except for the network for control decisions. The whole system is implemented on both of 16 bit personal computers and 32 bit workstations.

First, we briefly provide the research background and the inverted pendulum problem itself, followed by the description of our composite neural network model. Next, some results from the simulation are given, which are subsequently compared with the results from a control system based on modern control theory. Then, some discussions and conclusion follow.

2 Research Background

For our study, we have chosen the broom balancing problem. The broom balancing problem is comparatively easy to express in mathematical notation, and is famous for being an experimental example that can be used to show the effectiveness of modern control theory [2][3][4][5].

Although mathematically designed control systems may offer high quality control for control objects it was designed for, it will be useless for other control objects, and thus redesigning the system will be necessary. On the other hand, humans are so general that they have fair control of a wide range of control objects. One of the major reasons for this difference occurs from the difference in the knowledge representation methods. For example, Ichikawa proposed a general-purpose control system based on adaptive production systems and applied it to the broom balancing problem[1]. On the other hand, recently attention is being focused on neural networks. There is growing expectation for its application to control, especially since researchers for neural networks applied the back propagation algorithm to a variety of problem domains[6]. Also, the broom balancing problem is taken as an example domain for neural network applications[7].

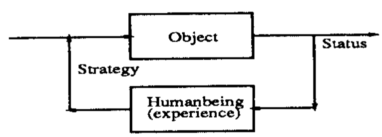

Fig.1: A human operator and controled object

Fig.1 shows a block diagram that represents a system with a controled object of control and a human controlling it. The network learns human experience by observing the system of Fig.1 from the outside (Fig.2).

Data equivalent to the input information needed to produce human control signals is regarded as the input signal for the neural network and the control signal produced by a human operator corresponding to the input information becomes the teacher signal. Neural networks learn by the back propagation. That is to say, the

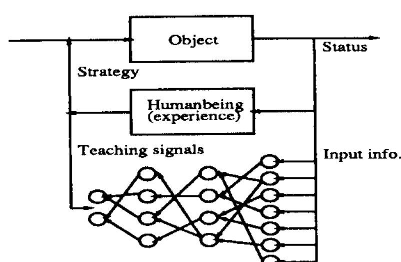

Fig.2: The neural network learns human operation

neural network system obtains the empirical knowledge of the human operator. After learning, the neural network is expected to be able to estimate the environment just as the human operator does(Fig.3). In this study, we will show that the neural network system mentioned above can be costructed.

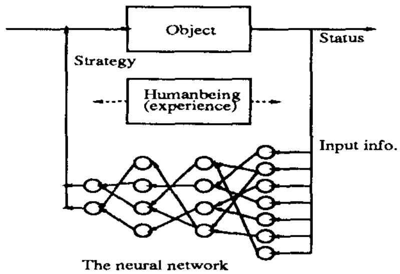

Fig.3: After learning

3 The Broom Balancing System

The broom balancing problem we deal with here is represented by the following system.

- The system consists of a cart and a pole that stands on the cart. The cart can move on a straight rail freely, and we will call this system, "The broom balancing system".
- The cart position can be controlled within ±5m, and the pole angle within ±24 degrees. Four different forces can be applied to the cart.

  LEFT STRONG, LEFT WEAK
  RIGHT STRONG, RIGHT WEAK
- The pole does not slip on the cart.
- Friction on the cart and air resistance on the pole are to be ignored.

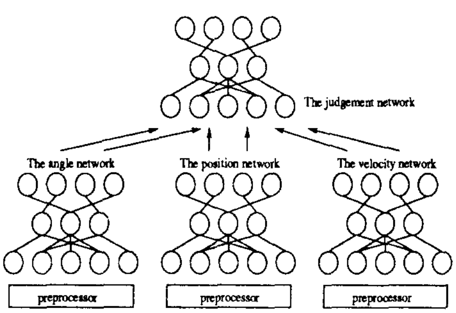

Fig.4: Overview of our composite neural networks system

4 Structure of the Neural Network

Fig.4 shows the overview of the neural network used in this study. The system is composed of three subnetworks - the angle network, the position network and the velocity network, which make up the criteria for determining the direction of output, and the judgement network, which integrates the output of the three subnetworks and makes the final estimation. We assume that there is a preprocessor for each subnetwork which converts visual pattern to input for each of the subnetworks. Each of the subnetwork corresponds the physical parameter - the angle of the pole, the position of the cart, and the cart velocity. Processing low level information is not the goal of this study. So we assume here that preprocessors for visual pattern recognition are already implemented. Although we will not use it here, Ralph Linker showed that, by utilizing neural networks, a preprocessor can be composed[9]. He used self-organization properties of neural networks to classify bitmap data. In the following section we describe each of the subnetworks.

4.1 The Three Subnetworks

The Angle Network The angle network makes estimation for angle information. It is a three-layered hierarchical neural network with one layer consisting of 20 input units, one with three hidden units, and one with four output units. We assume that the preprocessor provided at the input of the angle network converts visual input information into physical angular parameters and adjusts it for network input.

The network classifies the angle of the pole into the following four classes.

- If the pole leans to the right, then output is 0010.
- If the pole leans to the left, then output is 0010.
- If the pole leans over the right threshold angle, then output is 0001.
- If the pole leans over the left threshold angle, then output is 1000.

The Position Network The position network makes the estimations for cart position information. It is a three-layered hierarchical neural network with one layer consisting of 20 input units, one with ten hidden units, and one with four output units. We assume that a preprocessor that converts visual information into input for the network is provided. This network classifies the position of the cart according to the following threshold values:

- a) extremely left
- b) left
- c) right
- d) extremely right.

The Velocity Network The velocity network makes estimation for cart velocity information. It is a three-layered hierarchical neural network with one layer consisting of ten input units, one with four hidden units, and one with three output units. We assume that a preprocessor that converts visual information into input for the network is provided. This preprocessor is implemented by using optical flow. This network classifies the velocity of the cart according to the following threshold values:

- a) exceeding the threshold velocity in the left direction
- b) none
- c) exceeding the threshold velocity in the right direction.

4.2 The Judgement Network

Control signals are produced by inputing information from the three subnetworks into the judgement network. Therefore we prepared 4+4+3 = 11 input units, four hidden units and five output units. We have not focused yet on how to determine the number of hidden units. In this study, we determined the number of hidden units by trial and error based on results of past experiments with the system. The threshold values for the angles, positions, and velocities were also selected experimentally.

5 Learning Experiments

The broom balancing system simulator was implemented for collecting the teacher signals from records of human actions. A human operator produces control signals (which later become teacher signals) by observing the broom balancing system. The simulator was implemented on a personal computer with high graphic input-output capability. Current multiple task OS's do not have the ability to maintain constant speeds needed for our animation. We mainly used a high speed workstation for the learning procedure.

Production of teacher signals from records of human actions is an interesting aspect of our system. But at the same time, this system presents some problems. Even under the exactly same conditions (i.e. the pendulum is in the same angle, and the position and velocity of the cart are the same), a human operator does not always produce the same output. A human operator can produce many control signals in only 30 seconds of control action. This is not always preferable. A contradiction in data can disturb the learning process. The teacher signals were basically produced according to the following rules. For a state P (consisting of the angle of the pendulum, the velocity of the cart, and the position of the cart), let R_P be the teacher signal (i.e. the control signal) in the right-direction, and L_P be the signal in the left-direction. If either

$$\frac{R_P}{L_P} \gg 1 \tag{1}$$

or

$$\frac{L_P}{R_P} \gg 1 \tag{2}$$

is held, we chose the data with the greater number of appearance to be the teacher signal. For every other situation, possibly correct control signals were considered one by one and an appropriate one was selected to be the teacher signal.

In this study, each neural network learned independently. We made each subnetwork learn a threshold value. Since the teacher signals also become the input for the judgement network, we later determined teacher signals with which the system as a whole can produce preferable control signals by trial and error.

The initial values for the connection weights for all the neural networks were determined randomly. Biases of the networks are predetermined as one [6]. We adopted, as the number of hidden units, the number for the established set of teacher signal and input signal when the squared error between the output of the neural network and the teacher signal was very close to 0 (under 0.00001). Finally, by using a selected number of hidden units, the connection weights of the judgement network converged at approximately 900 cycles of the back propagation algorithm.

Table 1: Control performance by using the composite neural network

	mass of pole	length	controllable max angle
Learned task	0.1 Kg	1.0 m	9.0 (deg) [≡ 100 %]
Length changed	–	→ 3.0 m	6.0 (deg) [67 %]
	–	→ 0.5 m	9.0 (deg) [100 %]
Mass changed	→ 0.5 Kg	–	8.1 (deg) [90 %]
	→ 1.0 Kg	–	6.1 (deg) [68 %]
	→ 0.05 Kg	–	9.0 (deg) [100 %]
Length and Mass changed	→ 0.05 Kg	→ 1.8 m	8.0 (deg) [89 %]

6 Experiments for Performance Evaluation

An advantage of using neural networks for control is in the flexibility gained in the system. By using a composite neural network which has finished its learning process, we tried to control an object with different physical parameters (Table 1). The learning process was conducted with the values of the weight of the pendulum equal to 0.1Kg, with length 1.0m, and inital angle at 8 degrees. We considered the control performance to be a success if prespecified conditions are satisfied for a certain period of time. (In our case, we defined it to be 200 seconds.)

7 Comparison with modern control theory

We have also made broom balancing simulations based on modern control theory.

Since controlling the pendulum to remain perpendicular is necessary for the broom balancing problem, we used linear differential equations in the neighborhood of the perpendicular position.

Let m be the mass of the pendulum, I be the moment of inertia of the pendulum, M be the mass of the cart, H be the drag imposed on the fulcrum in the x direction and V be the drag in the y direction.

Then the equation of motion for the broom balancing system can be expressed as follows:

$$\left.\begin{array}{rcl} I\frac{d^2}{dt^2}\varphi & = & Vl\sin\varphi - Hl\cos\varphi \\ m\frac{d^2(l\cos\varphi)}{dt^2} & = & -mg + V \\ m\frac{d^2(x+l\sin\varphi)}{dt^2} & = & H \\ M\frac{d^2}{dt^2}x & = & u - H. \end{array}\right\} \quad (3)$$

If we assume $\varphi \approx 0$, and so $\sin\varphi \approx \varphi$, $\cos\varphi = 1$ in equation (3), and thus (3) becomes:

$$\left.\begin{array}{rcl} \frac{1}{3}ml^2\ddot{\varphi} & = & vl\varphi - hl \\ v & = & mg \\ m\ddot{x} + ml\ddot{\varphi} & = & h \\ m\ddot{x} & = & f - h. \end{array}\right\} \quad (4)$$

In modern control theory, linear approximations are often used to solve the problem in nonlinear regions. In other words, in modern control theory, control signals can be produced in regions where linear approximations are possible. However, it is difficult to produce signals in regions other than these.

In order to make comparisons with modern control theory, we chose the following four poles for computation:

$$\begin{array}{rclrcl} \delta_1 & = & -0.5 & \delta_2 & = & -1 \\ \delta_3 & = & -0.5 + j0.5 & \delta_4 & = & -0.5 - j0.5 \end{array} \quad (5)$$

The results of the simulations based on equation (3), (4) and (5) are shown in Table 2. Systems based on modern control theory is empirically better than our system in the linear area at stability and control cost (in this case, it means energy used by cart movement). On the other hand, the neural network system has flexibility. The neural network control system in this paper is able to control about 25% larger than systems based on modern control theory. The control signals generated by our system are similar to ones in bang-bang control. If we were to make comparisons in wider regions which extend out into the nonlinear regions and where physical parameters differ for each object, our system seem to be effective.

Table 2: Control performance based on modern control theory

	mass of pole	length	controllable max angle	
Learned task	0.1 Kg	1.0 m	5.8 (deg)	[≡ 100 %]
Length changed	–	→ 1.2 m	4.9 (deg)	[84 %]
	–	→ 1.5 m	2.6 (deg)	[45 %]
	–	→ 1.6 m	— (deg)	[0 %]
	–	→ 0.5 m	7.2 (deg)	[124 %]
Mass changed	→ 0.5 Kg	–	5.5 (deg)	[95 %]
	→ 1.0 Kg	–	4.8 (deg)	[83 %]
	→ 0.05 Kg	–	5.8 (deg)	[100 %]
Length and Mass changed	→ 0.05 Kg	→ 1.8 m	6.2 (deg)	[107 %]

8 Discussion

Before the subnetworks mentioned above were selected, we considered other candidates such as:

1. The angle velocity network
2. The acceleration network
3. The angle acceleration network
4. The inclination-direction detection network
5. The velocities-direction detection network
6. etc.

By assuming appropriate preprocessors, these subnetworks can learn the criteria for each situation. At first, it was thought that, with more subnetworks, control capability would be increased. However, it became clear that there were many obstacles against the progress of experiments. It is difficult to produce teacher signals for instance, but more than that it was found that control capability and the amount of network information were not always directly proportional. So having many subnetworks does not always contribute to the increase in control capability although it slows down convergence of the learning process for the judgement network. This may be because the judgement network used in our experimental system was too small in size. In spite of these current uncertainty, as a result of considering many systems with various learning data and subnetworks, we concluded that through experience the three subnetworks mentioned above were generally satisfactory for our study.

We also considered the following points for the organization of data.

- The broom balancing problem is symmetrical. So the data for the right side can be considered equivalent to the ones for the left side.
- The validity of giving examples of human failure as the teacher signals.

However, in our study, we excluded data produced by the operators and caused contradiction in the learning process. This seems to be a meaningful method for our approach. The order of feeding the data did not influence control capability. We did not give any example of human failure as the teacher signal.

9 Conclusion

This study was conducted to show the possibilities of using composite neural networks to construct a system for contolling a nonlinear object. By learning actual data selected from human trial and error, and without any other characterization, our system can acquire human knowledge for the control object. However, numerous considerations needed for teacher signals, and trial and error needed to determine the number of units in a network, are problems that cannot be avoided to make the learning process converge properly. In the present study, we selected teacher signals so that states of the broom balancing system and the control signals produced by the human operator can be expressed in terms of either equation (1) or equation (2).

Excluding complication accompanied by the production of teacher signals, and by letting each modularized neural network learn individually and then combining, the computational load on the network can be alleviated. The advantage is that the frequency of learning can be held low, and the relearning of the network can be easily accomplished.

Furthermore , we confirmed experimentally that our neural network control system exceeded in performance a model based on modern contol theory. The advantages of our system compared to those based on modern control theory may include:

1. It is not necessary to formulate system dynamics precisely.
2. It is not necessary to make accurate measurement of parameters.
3. It is able to control pendulums for wider regions of parameters.
4. It is more robust to variation of parameters.

However, neural networks only learn how to give an output for a given situation. We cannot determine if a given condition produces control signals to keep the broom upright, or at the optimum. Besides, since our system could only produce control signals which are like the ones in bang-bang control, that is, since there may be a limit in the man-machine interface needed for producing control signals, it may deteriorate modern control theory in some aspects. Thus, our next goal may be to construct a more capable control system by enlarging the scale of the judgement network.

References

[1] A Ichikawa,"Versatile Control by Means of Adaptive Production System," Proc. of SICE,Vol.19, No.1,pp.88-90,1981.

[2] D.G. Luenberger,"An Introduction to Observers," IEEE Trans. A. C., AC-16-6,pp.596-602,1971.

[3] T.Tsujisawa, S.Yogo and R. Tagawa," Control of a Second-Dimensional Inverted Pendulum," Proc. of SICE, Vol.19, No.12,pp.28-33,1983.

[4] W. Maletinsky, M.F.Senning and F.Wiederkehr, "Observer Based Control of a Double Pendulum," the IFAC 8th World Congress Preprint, XIII, Session 63,1981.

[5] W.R. Strurgeon and M.V. Loscutoff,"Application of Modal Control and Dynamic Observers to Control a Double Inverted Pendulum," JACC Preprint,pp. 857-865,1972.

[6] D.E.Rumelhart,G.E. Hinton, and R.J.Williams," Learning Internal Representations by Error Propergation," Parallel Distributed Processing, Vol. 1, D.E.Rumelhart, J.L.McClelland, and the PDP Research Group, Cambridge, MA : MIT Press,1986.

[7] Viral V. Tolat and Bernard Widrow,"An adaptive neural net controller with visual inputs," Neural Networks, Vol.1 ,Sup 1, 1988.

[8] Y.Miyata, " The learning and planning of action," ICS report,1988.

[9] Ralph Linker," Self-Organization in a Perceptual Network," COMPUTER,pp.105-117, IEEE,1988.