INTELLIGENT CONTROL BASED ON FUZZY LOGIC AND NEURAL NET THEORY

Chuen-Chien Lee

Department of Electrical Engineering and Computer Sciences
University of California
Berkeley, CA 94720

ABSTRACT

In the conception and design of intelligent systems, one promising direction involves the use of fuzzy logic and neural network theory to enhance such systems' capability to learn from experience and adapt to changes in an environment of uncertainty and imprecision. This paper explores an intelligent control scheme by integrating these multidisciplinary techniques. A self-learning system is proposed as an intelligent controller for dynamical processes, employing a control policy which evolves and improves automatically. One key component of the intelligent system is a fuzzy logic-based system which emulates human decision-making behavior. Another key component is cognitive neural models derived from animal learning theory, which stimulate memory association and learning behavior. It is shown that the system can solve a fairly difficult control learning problem. Simulation results demonstrate that improved learning performance can be achieved in relation to previously described systems employing bang-bang control. The proposed system is relatively insensitive to variations in the parameters of the system environment.

I. INTRODUCTION

During the past several years, a highly promising direction in the design of intelligent systems has emerged. More specifically, the direction in question involves the use of fuzzy logic and neural network theory to enhance the ability of intelligent systems that can learn from experience and adapt to changes in an environment of uncertainty and imprecision. This paper provides a brief introduction on a fuzzy logic-based system [16][17] and cognitive neural models [18][19], and explores an intelligent control system by integrating these multi-disciplinary techniques. The approach described here may be viewed as a step in the development of a better understanding of how to combine a fuzzy logic-based system with a neural network to achieve a significant learning / adaptive capability.

A. Why Fuzzy Logic Control?

There are many complex industrial processes which cannot be satisfactorily controlled by conventional methods due to modeling difficulties and unavailability of quantitative data regarding input-output relations. And yet, skilled human operators can control such systems quite successfully without having any quantitative models in mind. Furthermore, the operation of many man-machine systems requires the use of rules of thumb, intuition, and heuristics. All of these features are uncertain and imprecise and cannot be addressed adequately by conventional methods. As the increasing complexity and nonlinearity of control systems render conventional methods less effective, a rule-based system based on fuzzy logic becomes an increasingly attractive alternative.

In fact, during the past several years, rule-based controllers based on fuzzy logic [16][17] have emerged as one of the most active and fruitful areas for research in the application of fuzzy set theory [34]. Among the representative applications of fuzzy logic-based controllers are the subway system in the city of Sendai [33], container ship crane control [32], elevator control [4][30], nuclear reactor control [2][11], automobile transmission control [23], air conditioners [22], anti-lock break systems [24] and human-quality robot eyes [5]. Experience shows that a rule-based controller using fuzzy logic make it possible to emulate and even surpass the decision-making ability of a skilled human operator.

Although there is an extensive literature describing various fuzzy logic-based controllers using approximate reasoning, the acquisition of the

This research was supported in part by NASA Grant NCC-2-275 and Micro Grant 88-094.

rule base in such controllers is not as yet well understood. In past applications, fuzzy decision rules are either obtained from verbal expressions or observations of human operator control actions. Since domain experts and skilled operators do not structure their decision making in any formal way, the process of transferring expert knowledge into a usable knowledge base is tedious and unsystematic. Our research aims at the development of a better understanding of such problems, with a view to enhancing the potential of fuzzy logic-based controllers, which can operate effectively in an environment of uncertainty and imprecision.

One direction that is beginning to be explored is that of the conception and design of fuzzy systems which have the capability to learn from experience. In this context, a combination of techniques drawn from both fuzzy logic and neural network theory may provide a powerful tool for the design of intelligent systems which can emulate the decision-making ability of a skilled human operator and the ability to learn and adapt to changes in an environment of uncertainty and imprecision.

B. Why Cognitive Neural Models?

The theory of animal learning is inferred from observed behavior and constitutes carefully testified postulates regarding elemental processes of learning. Recent research into animal learning can be separated into two categories: the behavioral and neural substrates of learning, namely, the psychological and physiological levels of learning. One way to bridge such a gap is to postulate neural analogies of behavioral modification paradigms. Hebb's postulate [9] for synaptic plasticity was the first trial as a neural analogy of associative learning, which attempted to bridge psychology and neurophysiology. The theory of adaptive networks originated with [9] and continues to be influenced by plausible neural analogies of behavioral conditioning [6][12] [7][28][26][29][13][27][14][8][15].

Contemporary artificial neural networks are frequently referred to as connectionist models, parallel distributed processing (PDP) models, and adaptive / self-organizing networks. Basically, it is a complex system of neuron-like processing units that operate asynchronously but in parallel and whose function is determined by the network topology of connectivity. Artificial neural networks provide a new computational structure, a plausible approach for information processing because of its adaptivity / learning as well as massive parallelism.

Although new learning algorithms and VLSI technologies have recently provided strong impetus to neural network research, many problems still exist. Among them, the comprehensibility of neural networks, theoretical parsimony / enormous cost, and limited empirical successes are some of the major issues underlying the limitations of current neural networks. The learning behavior of such networks is difficult to understand, and the role of generic elements and subnetworks is unclear. Furthermore, most of these networks lack a theoretical foundation. The time and effort required to develop neural network architectures (network topology) and training is very high. Research has been directed in the main at "modeling applications", while relatively few "fielded applications" have emerged [3]. Most of such applications are restricted to pattern recognition, categorization, and realizations of associative memory. They are still toy research problems at the proof-of-concept stage. Among the few exceptions, the Adaptive Channel Equalizer (developed by Bernard Widrow) is perhaps the most commercially successful of all neural network applications to date. It is a single-neuron device used now in virtually all long-distance telephone systems to stabilize voice signals [3].

Klopf [13] has postulated that, "An intelligent system will have to build on a foundation that amounts to a highly detailed, immense microscopic knowledge base, a knowledge base that can be interfaced effectively with higher functional levels." From this perspective, a neural substrate could develop into the microscopic knowledge base. The macroscopic capabilities of intelligence could then be built on top of this. Given the limitations of current neural networks, a plausible scheme is to incorporate capabilities previously found on the macroscopic, network level into the microscopic, neuronal (single-neuron) level.

In this connection, we introduce cognitive single-neuron models that coincide with existing animal learning theory. Each proposed model provides a basis for understanding and explaining Pavlovian conditioning [25][20] and instrumental conditioning [20], respectively, which are the best understood animal learning processes. In particular, one model, an associative critic neuron, captures the predictive nature of Pavlovian conditioning, which is essential to the theory of adaptive / learning systems. Another model, an associative learning neuron, possesses the associative nature of instrumental conditioning, which stores in memory the temporal relationship between input and output.

C. Outline

The problem of learning via credit assignment [4] is described in Section II. The statement of the pole-balancing problem follows. This problem may be viewed as a canonical example of dynamic control. Some concepts from earlier related work are given in Section III. They serve as a basis for comparison of previous and proposed approaches. The proposed intelligent system is presented in Section IV. Here, a fuzzy logic-based controller is introduced, and a learning system with cognitive neural models is proposed. Computer simulation results are described in Section V. The paper closes with a concluding remark in Section VI.

II. A CASE STUDY: THE POLE BALANCING PROBLEM

In machine learning, the problem of learning to control physical dynamical systems has been, and remains, a challenging goal. In this context, the credit-assignment problem is often encountered in adaptive problem-solving systems, and is especially acute when evaluative feedback is delayed or infrequent. Basically, the credit-assignment problem, is to determine a strategy for assigning positive credit (reward) to desirable actions and negative credit (punishment) to undesirable actions in a way that would lead to the achievement of a specific goal. In what follows, we describe an approach to the building of an intelligent rule-based system that can learn to control a dynamical system without prior knowledge of its input-output relations.

Our approach focuses on a paradigmatic control problem - the pole-balancing problem - which has been the object of several studies in the literatures of control and neural networks. The pole balancing system is described as follows. A rigid pole is hinged to a cart, which is free to move on a one-dimensional track. The pole can rotate in the vertical plane of the track and the controller can apply an impulsive force of bounded magnitude to the cart at discrete time intervals. By balancing the pole, we mean that the pole never deviates by more than, say, 12 degrees, from the vertical. The equations of motion of the cart-pole system are not known to the controller, which implies that the cart-pole system is treated as a black box. What is known is a vector describing the cart-pole system's state at every time step. If the pole falls, it receives a failure signal. After a failure signal has been received, the system is reset to its initial state and a new attempt is made. On the basis of this evaluative feedback, the controller must develop its own control strategy and learn to balance the pole for as long as possible. Since a failure signal usually occurs only after a long sequence of individual control decisions, the sparsity of this signal makes the credit-assignment problem nontrivial.

III. PREVIOUS RELATED WORK

There are two noteworthy previous studies which have addressed the pole-balancing problem. The first is that of Michie and Chambers [21] in 1968. They constructed a program called BOXES that learned to balance the pole by applying two opposite constant forces. The second study is that of Barto, Sutton, and Anderson [1] in 1983, which used neuronlike adaptive elements to solve the same problem by using two constant forces. In general, both approaches can handle the credit-assignment problem that we mentioned. In both, the state space is partitioned into several non-overlapping regions and no symbolic reasoning techniques are employed. Both are limited to only two control actions: pushing the cart left or right with a force of fixed magnitude. The problem is thus one of bang-bang control.

In contrast to these approaches, we attempt to solve the problem through the use of symbolic problem-solving techniques, employing a fuzzy rule-based controller with approximate reasoning. Furthermore, a continuous control scheme is employed, namely, the controller can apply a force with a magnitude within [-10,+10] newtons. In this way, better performance of the controlled system may be achieved but the complexity of the problem is increased substantially. An overlapping partition of the state space forms a linguistic space. The overlapping partition enhances the speed of learning and robustness. We will have more to say about these issues at a later point.

IV. THE INTELLIGENT CONTROL SYSTEM

Experience shows that a fuzzy logic-based system using approximate reasoning [16][17] make it possible to emulate and even surpass the decision-making ability of a skilled human operator. And, neural network theory [3] provide a new computational structure, a plausible approach for information processing because of its adaptivity / learning as well as massive parallelism. In this connection, We developed an intelligent control scheme by integrating human decision-making and animal learning behavior employing fuzzy logic and neural network theory.

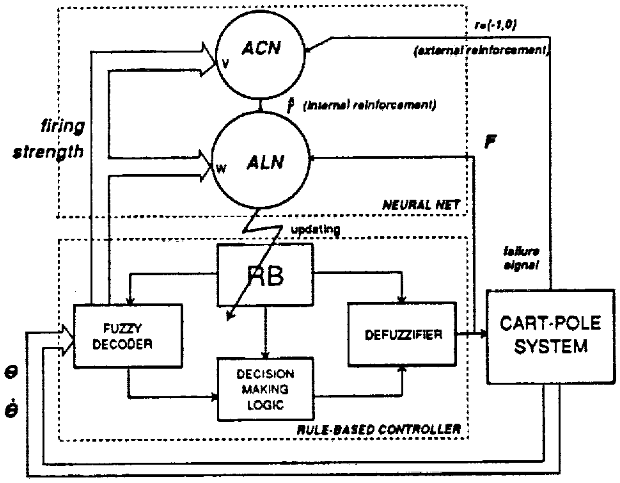

Fig. 1. Schematic representation of the intelligent system.

As shown in Figure 1, one key component of the intelligent system is a fuzzy logic-based controller which emulates human decision-making behavior. Another key component is a neural net. The net is composed by two cognitive neural models, an associative critic neuron (ACN) and an associative learning neuron (ALN), derived from animal learning theory, which stimulate memory association and learning behavior.

As a key component of the intelligent controller, the fuzzy logic-based system provides a linguistic description of control strategy. It is composed by a rule base, a fuzzy decoder, decision-making logic, and a defuzzifier. In general, the rule base describes control strategy which has the form of a collection of fuzzy control rules. For example, *if the angle of the pole is positive large and the angular velocity is positive large, then the applied force is positive large.* These are implemented and manipulated using fuzzy set theory [34] and are to be learnt by the proposed neural net. The fuzzy decoder inspects the incoming system state and fires the rules in parallel. A set of firing strength (x_i) is then generated and serves as input for the decision-making logic and neural net. The decision-making logic, the inference engine of the system, emulates human decision-making behavior based on the principles of approximate reasoning [35]. The defuzzifier takes a fuzzy control decision from the decision-making logic and determines a non-fuzzy control action (F).

The learning capability of the intelligent systems is provided by the associative critic neuron (ACN) and associative learning neuron (ALN). More specifically, the ACN is derived by using Pavlovian conditioning theory [25][20]. It captures the predictive nature of Pavlovian conditioning and has to do with criticism ($\hat{r}$) from the environment (r) associated with the system state (x_i). The ACN derives from the instrumental conditioning theory [20]. It is an associative memory system, which remembers the temporal relationships between input (x_i) and output (F), and associates each fuzzy control rule with an appropriate fuzzy control action (F_i).

A. Fuzzy Logic Control

In recent years, rule-based controllers employing approximate reasoning have emerged as one of the most active areas of research in the applications of fuzzy set theory. Such reasoning [35] plays an essential role in the remarkable human ability to make rational decisions in an environment of uncertainty and imprecision. In essence, approximate reasoning is the process or processes by which a possibly imprecise conclusion is deduced from a collection of imprecise premises. By employing the techniques of fuzzy set theory [34], approximate reasoning (with precise reasoning viewed as a limiting case) can be studied in a formal way.

The concept of a fuzzy set may be viewed as an extension of an ordinary (crisp) set. In a fuzzy set, an element can be a member of the set with a degree of membership varying between 0 and 1. Thus, a fuzzy set F in a universe $U = \{u_i, i=1, ..., n\}$ is defined by its membership function $\mu_F : U \rightarrow [0,1]$. If the $\mu_F(u_i)$ are 0 or 1, the fuzzy set is an ordinary set. As a special case, a fuzzy singleton is a fuzzy set containing just one element with degree 1.

A concept which plays an important role in the applications of the theory of fuzzy sets is that of a *linguistic variables.* To illustrate, if *speed* is interpreted as a linguistic variable, that is, a variable whose values are linguistic labels of fuzzy sets, then the values of *speed* might be

$$T(speed) = \{slow, moderate, fast,$$

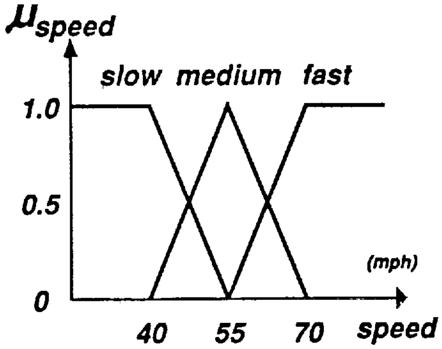

Fig. 2. Diagrammatic representation of various linguistic values of speed.

$$\textit{very slow}, \textit{more or less fast}, \cdots \}.$$

In a particular context, *slow* may be interpreted as, say, "a speed below about 40 mph", *moderate* as "a speed close to 55 mph" and *fast* as "a speed above about 70 mph". Figure 2 shows this interpretation in the framework of fuzzy sets.

The set-theoretic operations on fuzzy sets are defined via their membership functions. More specifically, let A and B be two fuzzy sets in U with membership functions μ_A and μ_B, respectively. The membership function $\mu_{A \cup B}$ of the *union* $A \cup B$ is defined pointwise for all $u \in U$ by

$$\mu_{A \cup B}(u) = \max\ \{\mu_A(u), \mu_B(u)\}.$$

Dually, the membership function $\mu_{A \cap B}$ of the *intersection* $A \cap B$ is defined pointwise for all $u \in U$ by

$$\mu_{A \cap B}(u) = \min\ \{\mu_A(u), \mu_B(u)\}.$$

If $A_1, \ldots, A_n$ are fuzzy sets in $U_1, \ldots, U_n$, respectively, the *Cartesian product* of $A_1, \ldots, A_n$ is a fuzzy set in the product space $U_1 \times \cdots \times U_n$ with the membership function

$$\mu_{A_1 \times \cdots \times A_n}(u_1, u_2, \cdots, u_n) = \min\ \{\mu_{A_1}(u_1), \cdots, \mu_{A_n}(u_n)\}.$$

Assume that the fuzzy sets A, A', B, and B' are the linguistic values of x and y. An example of approximate reasoning involving x and y is the following:

premise 1 : x *is* A',

premise 2 : *if* x *is* A *then* y *is* B,

consequent: y *is* B'.

For instance:

premise 1 : *the speed of a car is very high*,

premise 2 : *if the speed of a car is high*
then the probability of an accident is high,

consequent: *the probability of an accident*
is very high.

This type of fuzzy inference is based on the compositional rule of inference for approximate reasoning suggested by Zadeh [35].

A rule-based controller consists of a set of fuzzy control rules which are processed through the use of approximate reasoning. For simplicity, suppose that we have the two rules:

R_1 : *if* x *is* A_1 *and* y *is* B_1 *then* z *is* C_1,

or

R_2 : *if* x *is* A_2 *and* y *is* B_2 *then* z *is* C_2.

Approximate reasoning, given (x is A') and (y is B'), computes the degree of partial match between the user-supplied facts and the knowledge rule base as follows.
The degrees of match of (A_i and A) and (B_i and B) are given respectively by

$$\alpha_i = \max_u \min\{\mu_{A_i}(u), \mu_A(u)\},$$

$$\beta_i = \max_v \min\{\mu_{B_i}(v), \mu_B(v)\}.$$

The firing strength of the i^{th} rule is given by

$$x_i = \min\{\alpha_i, \beta_i\}.$$

Hence, the i^{th} rule recommends a control decision as follows:

$$\mu_{C_i'}(w) = \min\{x_i, \mu_{C_i}(w)\}.$$

The consequences of multiple rules can be combined by a conflict-resolution process which treats the sentence connective *or* as a union operator. The combined consequence is then given by

$$\mu_C(w) = \max\{\mu_{C_1'}, \mu_{C_2'}\}.$$

The combination of consequences is illustrated in Figure 3.

In on-line processes, the states of a control system are essential to a control decision (action). The underlying data are usually obtained from sensors and are crisp. It may be necessary to convert these data into the form of fuzzy sets [16]. In practice, however, crisp data are frequently treated as

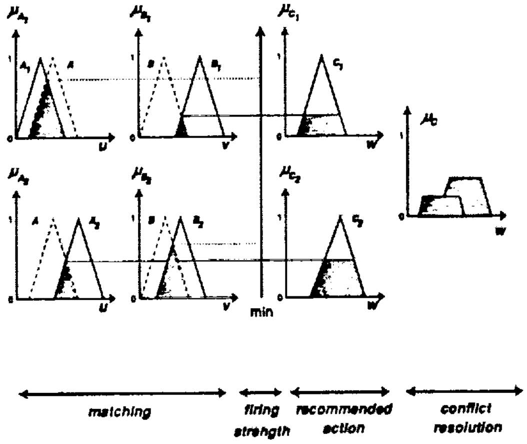

Fig. 3. Diagrammatic representation of approximate reasoning using fuzzy input.

fuzzy singletons. In this case, the corresponding inference mechanism is shown in Figure 4.

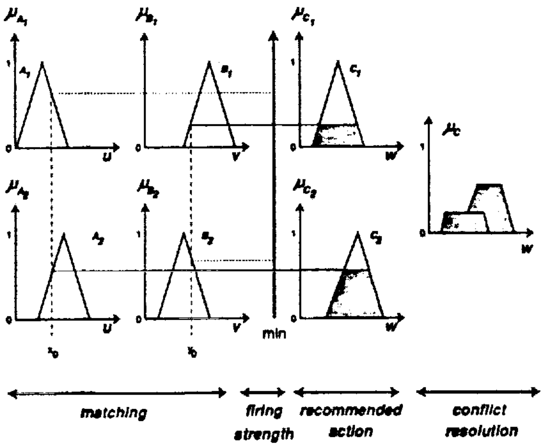

Fig. 4. Diagrammatic representation of approximate reasoning using crisp input.

Furthermore, in on-line control, the inference process should lead to a non-fuzzy control action. This necessitates the use of a *defuzzifier*. A defuzzifier can be implemented by using *max criterion, mean of maximum* or *center of area* algorithms [17]. The defuzzifier used here is employing the *center of area* algorithm.

In what follows, the fuzzy control rules are assumed to be of the form

$$R_i : \textit{if } x \textit{ is } A_i \textit{ and } y \textit{ is } B_i \textit{ then } z \textit{ is } C_i ,$$
$$i = 1, 2,..., n \ ,$$

where x, y, and z are linguistic variables representing the angle of the pole with respect to the vertical axis, angular velocity of the pole, and applied force, respectively; A_i, B_i, and C_i are the linguistic values (fuzzy sets) of the linguistic variables x, y, and z in their respective universes of discourse, [-12,+12] degrees, R, and [-10, +10] newtons. The definitions of linguistic values A_i and B_i are shown in Figure 5 (a) and (b). The problem is to learn the linguistic values C_i, which take the form of triangles, defined on the control force universe [-10,+10] newtons. The conception of fuzzification is performed as shown in Figure 5 (c). The location of the vertex of such a triangle is to be learned, while the coordinates of the base are functions of the vertex location value, say in the extreme case, +/-2 newtons away from that vertex.

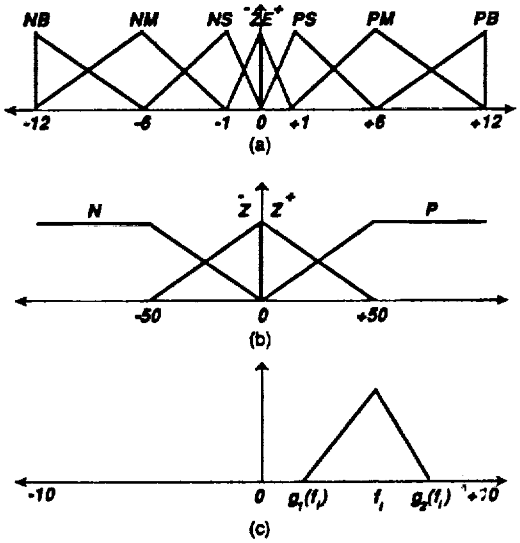

Fig. 5. (a) Linguistic values of angle, (b) angular velocity, and (c) applied force.

To summarize the ideas thus far discussed, the conception of a 2-D linguistic state space is formed. The x axis is θ with seven linguistic values; the y axis is $\dot{\theta}$ with three linguistic values. Thus, 8x4 fuzzy control rules are involved. Each fuzzy control rule corresponds to a fuzzy cell. The premise of a fuzzy control rule determines the cell's coordinates in the linguistic state space. The consequent of the rule is taken to be the content of the cell, which is to be learned by the proposed neurons, the ALN and ACN. Once a system input is sensed, the cells are fired in parallel. The fuzzy decoder takes the current state of the cart-pole system as an input and has n outputs (firing strengths) going to the ALN and ACN. Each output of the fuzzy decoder corresponds to a fuzzy cell. The activity of the output is the firing strength. The firing strength serves as an input to both the ALN and ACN, and is also used to compute the recommended control action in each rule (cell).

B. Learning with a Neural Net

As has been mentioned in Section II, the principal difficulty in the learning process is that the training information (failure signal) is very sparse. Many of the previously employed neural networks such as the Adaline, perceptrons, and Hopfield nets, are effective for the solution of supervised pattern classification problems. In contrast, our network consists of the ACN and ALN which perform unsupervised learning. The ACN has to do with the criticism from the environment associated with the system state. The ALN takes the criticism and associates n fuzzy control actions with n fuzzy cells (the consequents of n fuzzy control rules). Since the ACN predicts the criticism at every time step, the ALN can continuously update itself before the failure signal occurs. This is the basis for the solution of the credit-assignment problem.

1. ACN

The ACN is derived from Pavlovian conditioning theory [25][20]. The best known example of Pavlovian conditioning comes from Pavlov's research on the conditioned reflex of salivation by dogs. Prior to conditioning, when a dog hears the sound of a bell, it pricks its ears. Then, when the food is presented to it, it salivates. If this sequence of events is repeated, the dog soon starts to salivate in reaction to the sound of the bell. In effect, the dog has been "conditioned" to react to the bell. As can be seen, the sound of a bell can be used to predict the occurrence of salivation before the presence of food. This predictive relationship between food and the sound of a bell has important implications. Thus, the ACN captures this predictive nature of the Pavlovian conditioning.

The correspondence between Pavlovian conditioning and the behavior of our system is as follows. Food corresponds to the evaluative feedback (failure signal). The salivation by reflex is equivalent to an *external reinforcement* $r(t)$ with the value -1.0 if failure signal occurs, otherwise 0.0. The sound of a bell relates to the i^{th} fired fuzzy cell (fuzzy control rule) with firing strength x_i. The salivation resulting from the bell's sound is the *predictive reinforcement* $v_i(t)$ of the i^{th} fuzzy cell. It is worth noting that, in the extreme, the i^{th} rule with firing strength either 1.0 or 0.0 is the exact case of presence or absence of a bell's sound in the conditioning of a Pavlov dog. In other words, our ACN operates in a continuous mode, which treats Pavlovian conditioning as a special case. In effect, the ACN attempts to predict the reinforcement $v_i(t)$ that can eventually be obtained from the environment by choosing a control action for that fuzzy cell.

As an extension of single-input/single-output analogy, multiple inputs in the ACN necessitate an output which is a weighted sum of the predictive reinforcements of all fired fuzzy cells. The weighted sum $p(t)$ is the *total reinforcement* of all fired fuzzy cells at time t. Furthermore, an internal reinforcement $\hat{r}(t)$, the criticism, is generated as a temporal difference of the total predictive reinforcements.

As shown in Figure 1, the ACN has an external reinforcement input, $r(t)$, from the cart-pole system, n inputs, $x_i(t)$, $i=1, \ldots, n$, from corresponding fuzzy cells, and an output, $\hat{r}(t)$, as internal reinforcement signal (criticism) for the ALN and itself. The total reinforcement at time t is given by

$$p(t) = G(\sum_{i=1}^{n} v_i(t)x_i(t)),$$

where G could be a sigmoid-shaped function, identity function, mean of maximum algorithm or center of area algorithm. The associative learning rule for the i^{th} fuzzy cell is in part characterized by a local memory trace $\bar{x}_i(t)$ and the internal reinforcement $\hat{r}(t)$. The predictive reinforcement $v_i(t)$ of the i^{th} fuzzy cell (fuzzy control rule, fuzzy system state) is updated by

$$v_i(t+1) = v_i(t) + \beta\hat{r}(t)\bar{x}_i(t),$$

where β is a positive learning-rate parameter. The local memory trace is defined by

$$\bar{x}_i(t+1) = \lambda\bar{x}_i(t) + (1-\lambda)|x_i(t)v_i(t)|,$$

where λ, $0 \leq \lambda < 1$, is a trace-delay parameter. The trace takes the form of an exponential curve. It is strengthened by the degree of firing strength of the i^{th} fuzzy cell (fuzzy control rule) together with its current weight, and weakened if the rule is not fired. The trace thus keeps track of how long ago the i^{th} fuzzy control rule fired and also how often it was fired. The internal reinforcement is calculated as

$$\hat{r}(t) = r(t) + \gamma p(t) - p(t-1),$$

where γ, $0 \leq \gamma < 1$, is a discount-rate parameter. The internal reinforcement serves as criticism, depending on a relative difference of $p(t)$ and $p(t-1)$. If the pole does not fall and $\gamma p(t) > p(t-1)$, then $r(t)=0$ and $\hat{r}(t)>0$, a reward is given. If the pole does not fall and $\gamma p(t) < p(t-1)$, then $r(t)=0$ and $\hat{r}(t)<0$, and a punishment is effected. The discount factor γ implies a bias for the condition in which $p(t)$ equals $p(t-1)$. More specifically, once the pole does not fall and keeps in the same state, a reward is given through the use of a discount factor. On the other hand, if the pole falls, then $p(t)=0$, $r(t)=-1$ and $\hat{r}(t)<0$, and a punishment is issued. If $p(t-1)$ fully predicts the occurrence of the failure, there is no punishment. As shown, a negative feedback mechanism is implicitly incorporated into the internal reinforcement.

The proposed ACN model might be viewed as an extension of the Sutton-Barto model [18]. More specifically, in the context of animal learning phenomena, a sigmoid-shaped acquisition curve is observed. This is not simulated in the Sutton-Barto model. In our model, it can be achieved by making a change in the associative strength proportional to the current associative strength [18]. It has been demonstrated by computer simulation that the ACN accounts for many phenomena observed in Pavlovian conditioning, such as a sigmoid-shaped acquisition curve, inter-stimulus interval effects, trace conditioning, and delay conditioning. A more detailed discussion of this aspect of our model is described elsewhere [18].

2. ALN

The ALN is derived from the instrumental conditioning theory [20]. A simple example is teaching a dog to perform a trick. During training, if the dog does well, it is given a reward. If not, it is punished. After training, the dog has learned a new trick. The association of the dog's response and reinforcement has in effect been "conditioned". The correspondence between this conditioning and the ALN is as follows. A dog corresponds to the i^{th} fuzzy control rule with firing strength x_i. The response of the dog relates to the control force $(w_i f_i)$ of the i^{th} rule. The reinforcement as reward/punishment is equivalent to the internal reinforcement from the ACN. The ALN does the following: the i^{th} fuzzy control rule can produce correct control force of the i^{th} rule under the internal reinforcement from the ACN. In effect, the ALN is a content-addressable memory system which associates each fuzzy control rule with an appropriate fuzzy control action.

As shown in Figure 1, the ALN has an internal reinforcement input, $\hat{r}(t)$, from the ACN, n inputs, $x_i(t)$, $i=1, ..., n$, from the fuzzy decoder, a control action input, $F(t)$, from the defuzzifer, and n associative weights w_i $i=1, ..., n$, as outputs for the rule base. Each associative weight $w_i(t)$ is transformed -- by using the concepts of dynamical normalization and fuzzification -- into a fuzzy set having the form of a triangle as described in the previous section. Symbolically,

$$F_i(t) = fuzzifier\ (f_i(t)),$$

where $f_i(t)$ is the location of the vertex of the triangle. It is given by

$$f_i(t) = H(w_i(t) + noise(t)), \quad i = 1, ..., n,$$

where H is a dynamic sigmoid function which may be viewed as a dynamic normalization function and provides a continuous output within the range [-10,+10]. For the purpose of computer simulation, the following function is used:

$$H(x,t) = \begin{cases} \dfrac{10x}{T(t)+x} & x>0, \\ \varepsilon & x=0, \\ \dfrac{10x}{T(t)-x} & x<0. \end{cases}$$

where $T(t) = k_1 \max\limits_i |w_i(t)|$ is an offset-tuning parameter which determines the slope of the sigmoid-shaped curve; and k_1 is a constant. The associative learning rule for each $w_i(t)$ is

$$w_i(t+1) = w_i(t) + \overline{\alpha}(t)\hat{r}(t)e_i(t),$$

$$\overline{\alpha}(t) = \frac{\alpha k_2}{k_2 + t},$$

where $\overline{\alpha}$ is a dynamical positive learning-rate parameter with a initial value α and k_2 is a weight-freeze parameter. The weight-freeze parameter determines the decreasing rate of the dynamical learning rate $\overline{\alpha}$. $\hat{r}(t)$ is the criticism from the ACN. The associativity trace, $e_i(t)$, is given by

$$e_i(t+1) = \delta e_i(t) + (1-\delta)F(t)x_i(t),$$

where δ, $0 \leq \delta < 1$, is another trace-decay parameter. The associativity trace takes the form of an exponential and it remembers for how long and how often a fuzzy control rule has fired as well as what control action was taken at that time.

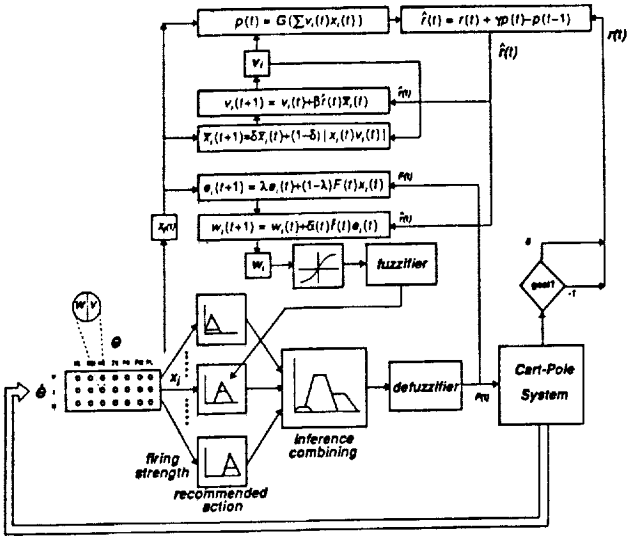
$p(t) = G(\sum v_i(t)x_i(t))$
$\hat{r}(t) = r(t) + \gamma p(t) - p(t-1)$
$r(t)$
$\hat{r}(t)$
v_i
$v_i(t+1) = v_i(t) + \beta\hat{r}(t)\bar{x}_i(t)$
$\bar{x}_i(t+1) = \delta\bar{x}_i(t) + (1-\delta)\,|x_i(t)v_i(t)|$
$e_i(t+1) = \lambda e_i(t) + (1-\lambda)F(t)x_i(t)$
$F(t)$
$w_i(t+1) = w_i(t) + \delta(t)\hat{r}(t)e_i(t)$
$x_i(t)$
w_i
fuzzifier
goal?
0
-1
w | v
θ
$\dot{\theta}$
x_j
firing strength
recommended action
inference combining
defuzzifier
Cart-Pole System

Fig. 6. Signal flow of the intelligent control system.

Figure 6 illustrates the signal flow of our proposed controller during a learning process. In principle, once a system state is sensed, the set of fuzzy control rules is fired in parallel. A set of firing strengths (x_i) is then generated and serves as input to both ALN and ACN. The information about the system state is then fed into the two neuronlike elements by the set of firing strengths. The firing strength together with the predictive reinforcement (v_i, or desirability) of the i^{th} fuzzy rule generates the local memory trace ($\bar{x}_i$, desirability trace) of the i^{th} fuzzy rule. The total reinforcement, p, or equivalently, the desirability of all fired fuzzy cells, is computed based on the firing strength and the reinforcement (desirability) of each fuzzy rule. A non-fuzzy control action, F, is determined after the processes of inference combining and defuzzification. The control action, F, together with the firing strength, x_i, of each rule contributes the associativity trace, e_i, of each rule. After applying the control action to the plant, a goal evaluation, r, is made, which takes binary values. Based on the yes-no evaluation, the criticism, $\hat{r}$, which is a more informative evaluation, is generated. It plays an important role in the solution of the credit-assignment problem. The weights (v_i, w_i) in learning rules are thus updated on the basis of the criticism and their own local memory trace, ($\bar{x}_i$, e_i). A fuzzy control force in each rule is generated from the w_i by the use of dynamic normalization and fuzzification.

V. SIMULATION RESULTS

We implemented our system on a Sun workstation. For comparison purposes, we also implemented Barto's system [1] for solving the same problem. The mass of the cart and initial pole were 1.0 kg and 0.1 kg, respectively. The length of the pole was 1.0 meter. The parameter values used in our simulation were: α=1000, β=0.5, γ=0.95, δ=0.9, λ=0.8, ε=0.1, k_1=0.15, and k_2=2500. A run was called "success" whenever the number of steps before failure was greater than 60,000. The external reinforcement $r(t)$ was -1 when the failure signal occurred, otherwise, it was 0. Every trial began with the same initial cart-pole states, θ=0, $\dot{\theta}$=0, x=0, $\dot{x}$=0, and ended with a failure signal when $|\theta|$>12 degrees. All memory traces, x_i and e_i, were set to zero. All the weights, w_i; were set to zero, and a lower bound v_i (=-0.0001) was set to all the weights. In testing the performance of the system, the simulator was run by applying the Adams predictor-corrector method with a time step of 20 ms [19].

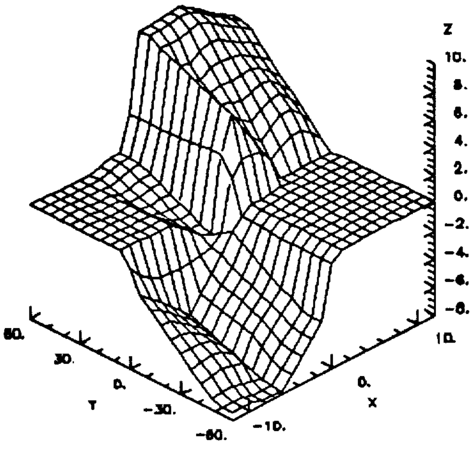

Fig. 7. Learned control surface based on the proposed intelligent system with COA defuzzifier.

A. Learning / Training

The proposed controller and Barto's system are capable of learning to balance the pole. However, experiments show that our system has a better learning performance [19]. The proposed controller learns to balance the pole by 6 trials with COA defuzzifier. Figure 7 illustrates the learned control surfaces based on our intelligent system employing defuzzifier *center of area*(COA). The performance of Barto system, in average, took 27 trials to balance the pole [19].

Additional observations were made on the state trajectory of the angle of the pole with respect to the vertical axis. We observed the data after the systems learned their own control strategy. The data showed that, in every case, our controller could keep the angle within a smaller region compared with Barto's. Figure 8 illustrates one set of these data from our system and Barto's, respectively.

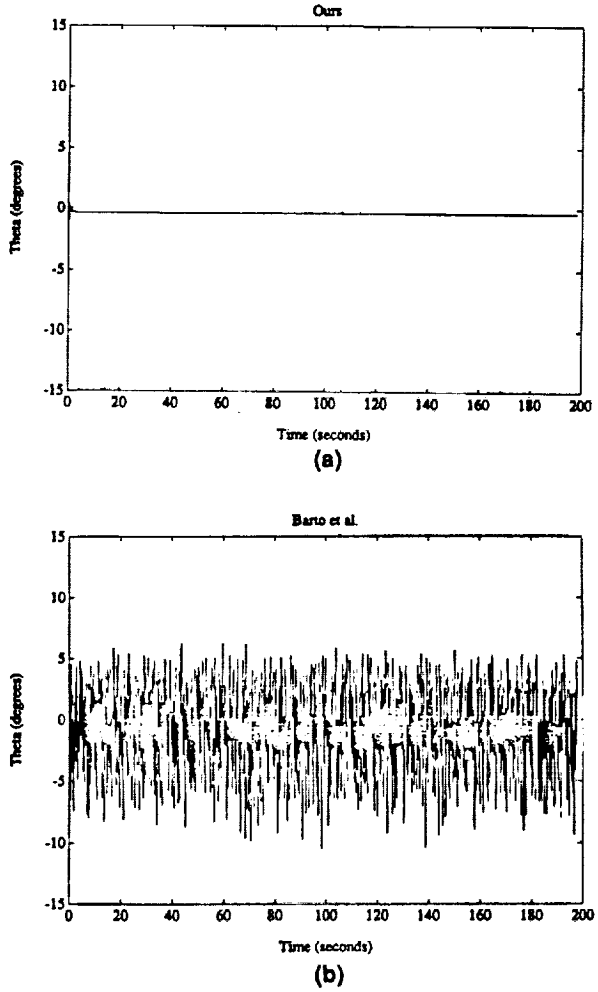

Fig. 8. (a) State performance of the pole angle based on the proposed controller. (b) State performance of the pole angle based on Barto's system.

B. Adaptation

Adaptation is intended to adjust to unforeseen changes in environmental conditions using prior knowledge. Training involves constructing a knowledge base of an application domain (e.g. a pole-balancing task) with little a priori domain knowledge. The capability of learning to solve new tasks by modifying previous learned knowledge (adaptation) is compared with that of starting from scratch (training). Extensive simulation studies of such schemes have been carried out. They show that the proposed controller tolerates a wide range of uncertainty as well as a lack of system information, e.g., parameter changes in the length and mass of the pole, changes of failure criteria, and a slanted cart-pole system.

The adaptation experiments were based on pre-learned knowledge by employing the same parameter settings as that in the last section. The length and mass of the pole were 0.1kg and 1.0m, the angle constraint for failure evaluation was -/+12°, and the initial value of the angle of the pole with respect to the vertical axis is 0.0°. The system took 6 trials to learn the task.

In the first set of experiments, the system is required to adapt to changes in the length and mass of the pole. Six experiments were performed. The first two were to increase the original mass of the pole by a factor of 10 and 20, respectively. The third and fourth ones were to change the original length of the pole by a factor of 2 and 1/2, respectively. The last two were to replace the original pole by two shorter poles. The length and mass of the first pole were reduced to 2/3 of the original values, while the second one is 1/4. Without pretraining, the system took 10, 15, 5, 11, 8 and 6 trials to learn these tasks. However, with the pre-trained knowledge, the system successfully completed these tasks without any further trials. The result shows the robustness of the proposed intelligent system.

In the second set, we added a more severe constraint on the angle of the pole for failure evaluation. The angle constraints were changed from +/-12° to +/-6°, +/-3°, and to +/-1°, respectively. The system needed 4 and 6 trials to learn the first two tasks with no initial knowledge, but it failed in the last task since a finer partition of input space is required. While with pre-training, the system adapted to all tasks without further trials.

In the third set, the system was required to adapt to the changes in the length and mass of the pole (by a factor of 1/2) and angle constraint (+/-3°). The training took 6 trials, while adaptation can handle the new task well.

Finally, the cart-pole system was lifted at the right end in such a way that the base of the system and the surface of the table formed an angle of 12°. The system took 10 trials to balance the pole. However, the system with the trained knowledge needed no further trials to complete the new task.

VI. CONCLUDING REMARK

In this article, we have proposed a symbolic problem-solving approach to a class of learning control problems. More specifically, we have attempted to develop an intelligent control scheme by integrating human decision-making with a fuzzy logic-based system and animal learning behavior with cognitive neural models. The proposed intelligent control system learns and improves its rule base for better control strategy from experience and adapts to changes in an environment of uncertainty and imprecision. In this way, we avoid an ad-hoc rule-tuning process which is usually inefficient and lacking in consistency. It has been shown that the proposed intelligent system has a better performance of learning speed and system behavior in relation to previous approaches. Furthermore, the system is quite robust. The controller is relatively insensitive to variations in the parameters of the system environment, e.g., in the context of pole-balancing, changes in the length and mass of the pole, failure criteria, and slanting the base of the cart-pole system. In addition, the controller can be primed with pre-trained control knowledge which minimizes rapid changes during adaptation.

The approach described in this paper may be viewed as a step in the development of a better understanding of how to combine a fuzzy logic based system with a neural network to achieve a significant learning capability. We plan to address various aspects of this important issue in subsequent papers.

ACKNOWLEDGMENT

I am greatly indebted to Professor Lotfi A. Zadeh of the University of California, Berkeley for his encouragement of this research. The assistance of Professor Zadeh is gratefully acknowledged.

REFERENCES

1. A. G. Barto, R. S. Sutton, and C. W. Anderson, "Neuronlike adaptive elements that can solve difficult learning control problems", *IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-13, no. 5,* pp. 834-846, 1983.
2. J. A. Bernard, "Use of a rule-based system for process control", *IEEE Control Systems Magazine, vol. 8, no. 5,* pp. 3-13, Oct. 1988.
3. *DARPA Neural Network Study,* Fairfax: AFCEA International Press, Dec. 1988.
4. Fujitec, "FLEX-8800 series elevator group control system," *Fujitec Co., Ltd,* Osaka, Japan, 1988.
5. Fujitsu, "Fujitsu Announces Creation of Human-Quality Robot Eye," *The Japan Times* Jan. 10, 1990.
6. S. Grossberg, "Some networks that can learn, remember, and reproduce any number of complicated space-time patterns," *Journal of Mathematics and Mechanics vol. 19,* pp. 53-91, 1969.
7. S. Grossberg, "Classical and instrumental learning by neural networks," in *Progress in Theoretical Biology vol. 3,* R. Rosen and F. Snell, Eds., New York: Academic Press, pp. 51-141, 1974.
8. S. Grossberg and N. A. Schmajuk, "Neural Dynamics of attentionally-modulated Pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing," *Psychobiology,* pp. 195-240, 1987.
9. D. O. Hebb, *The Organization of Behavior,* New York: Wiley, 1949.
10. Y. Kasai and Y. Morimoto, "Electronically controlled continuously variable transmission," *Proc. Int. Congress on Transportation Electronics,* Dearborn, Michigan, Oct. 1988.
11. M. Kinoshita, and T. Fukuzaki, T. Satoh, and M. Miyake, "An automatic operation method for control rods in BWR plants," *Proc. Specialists' Meeting on In-core Instrumentation and Reactor Core Assessment,* Cadarache, France, 1988.
12. A. H. Klopf, "Brain function and adaptive systems - A heterostatic theory," *Aire Force Cambridge Research Laboratories Special Report AFCRL-72-0164,* 1972.
13. A. H. Klopf, *The Hedonistic Neuron,* Washington: Hemisphere, 1982.
14. A. H. Klopf, "A drive-reinforcement model of single neuron function: An alternative to the Hebbian neuronal model," in *Neural Networks*

for Computing, Snowbird, Ut., 1986 J. S. Denker, Ed., New York: American Institute of Physics, 1986, pp. 265-270.

15. A. H. Klopf, "Drive-reinforcement learning: A real time learning mechanism for unsupervised learning," in *IEEE 1st International Conference on Neural Networks,* San Diego, June, 1987.

16. C. C. Lee, "Fuzzy logic in control systems: Fuzzy logic controller - Part I," *IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-20, no. 2,* 1990.

17. C. C. Lee, "Fuzzy logic in control systems: Fuzzy logic controller - Part II," *IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-20, no. 2,* 1990.

18. C. C. Lee, "Modeling behavioral substrates of associative learning and memory: adaptive neural models," *IEEE Trans. on Systems, Man, and Cybernetics,* to appear.

19. C. C. Lee, "A Self-Learning Rule-Based Controller Employing Approximate Reasoning and Neural Network Concepts," *International Journal of Intelligent Systems, vol. 5, no. 3,* 1990.

20. N. J. Mackintosh, *The Psychology of Animal Learning,* New York: Academic Press, 1974.

21. D. Michie and R. A. Chambers, "Boxes: an experiment in adaptive control," *in Machine Intelligence 2,* E. Dale and D. Michie, Eds., Oliver and Boyd, Edinburgh, 1968, pp. 137-152.

22. Mitsubishi, "World's First Fuzzy Logic Controlled Commercial Air Conditioner Begins Volume Shipping," *NEWS - Togai InfraLogic Inc.* Dec. 19, 1989.

23. Nissan, "New auto systems use fuzzy logic," *New York Times,* July 1, 1989.

24. Nissan, "Nissan Patents 'Fuzzy Logic' ABS, Gearbox," *Automotive Electronic News,* July 31, 1989.

25. I. P. Pavlov, *Conditioned reflexes,* London: Oxford University Press, 1927.

26. J. M. Pearce and G. Hall, "A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli," *Psychological Review vol. 87, no. 6,* pp. 532-552, 1980.

27. J. M. Pearce, H. Kaye, and G. Hall, "Predictive accuracy and stimulus associability: Development of a Model for Pavlovian learning," in *Quantitative Analyses of Behavior III,* M. L. Commons, R. J. Herrnstein, and A. R. Wagner, Eds., New York: Harper and Row, Publishers, pp. 241-255, 1982.

28. R. A. Rescorla and A. R. Wagner, "A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement," in *Classical Conditioning II: Current Research and Theory,* A. H. Black and W. F. Prokasy, Eds., New York: Appleton-Century-Crofts, pp. 64-99, 1972.

29. R. S. Sutton and A. G. Barto, "Toward a modern theory of adaptive networks: expectation and prediction," *Psychological Review vol. 88, no. 2,* pp. 135-170, 1981.

30. Toshiba, "'Command-AI' Elevator Systems," *Nikkei Industry News,* Sept. 2, 1989.

31. A. R. Wagner, "Expectancies and the priming of STM," in *Cognitive Processes in Animal Behavior* S. H. Hulse, H. Fowle, and W. K. Honig, Eds., New York: Halsted Press, pp. 177-209, 1978.

32. S. Yasunobu and G. Hasegawa, "Evaluation of an automatic container crane operation system based on predictive fuzzy control," *Control Theory and Advanced Technology, vol. 2, no. 3,* pp. 419-432, 1986.

33. S. Yasunobu and S. Miyamoto, "Automatic train operation by predictive fuzzy control," in *Industrial Applications of Fuzzy Control,* M. Sugeno, ed., Amsterdam: North Holland, 1985.

34. L. A. Zadeh, "Fuzzy sets," *Information and Control, vol. 8,* pp. 338-353, 1965.

35. L. A. Zadeh, "Outline of a new approach to the analysis complex systems and decision processes," *IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-3,* pp. 28-44, 1973.

An Neural Network Based Fuzzy Logic Controller

(Paper not provided by publication date.)