

# FUZZY CONTROL AS A UNIVERSAL CONTROL TOOL

Hung T. Nguyen<sup>1</sup>, Vladik Kreinovich<sup>2</sup>, Ongard Sirisaengtaksin<sup>3</sup>

<sup>1</sup>*Department of Mathematical Sciences,  
New Mexico State University, Las Cruces, NM 88003*

<sup>2</sup>*Computer Science Department,  
University of Texas at El Paso, El Paso, TX 79968*

<sup>3</sup>*Department of Computer and Mathematical Sciences,  
University of Houston-Downtown, Houston, TX 77002*

**Abstract.** It is known that fuzzy control is a *universal control tool*, because an arbitrary control strategy (in particular, a control strategy that is in some sense optimal) can be obtained in principle by applying a fuzzy control methodology to some set of rules. This result has already been proved (e.g., [2,7,20,21]) for the case when a plant is described by finitely many parameters and a special type of fuzzy control methodology is used. In this paper, we prove it for arbitrary plants (including plants that are *distributed systems*, i.e., plants whose state requires infinitely many parameters to describe) and arbitrary fuzzy control methodologies. We also prove that there exists a *universal fuzzy controller* that generates an appropriate control from an input description of a plant (and the desired objective). Mathematically, we prove that fuzzy systems can approximate arbitrary continuous functionals, thus generalizing a known result about continuous functions.

**Keywords:** control theory, universal approximator, distributed systems, universal controller.

## 1. INTRODUCTION

**Traditional control methods are not always applicable.** Assume that we have to control a plant. If we know exactly how this plant works and how it reacts to different controls, if we know exactly what we want from the control, and if the corresponding equations are not extremely complicated, then methods of traditional control theory usually provide us with an optimal (or at least reasonably good) control strategy. In many real-life situations, however, we do not have complete knowledge of the plant; in such situations, we cannot apply the traditional control methodology.

**Fuzzy control methodology.** In situations where we do not have complete knowledge of the plant, we often have the experience of human operators who successfully control this plant. We would like to make an automated controller that uses their experience. With this goal in mind, an ideal situation is when an operator can describe his control strategy in precise mathematical terms. Most frequently, however, the operators cannot produce such a description (can you describe how exactly you drive your car?); instead, they explain their control in terms of *rules* formulated in natural language (like “if the velocity is high and the obstacle is close, brake immediately”). *Fuzzy control* is a methodology that translates these natural-language rules into an automated control strategy. This methodology was first outlined by Zadeh [3] and experimentally tested by Mamdani [13] in the framework of *fuzzy set theory* [22] (hence the name). For many practical systems, this approach

works just fine (for the current state of fuzzy control the reader is referred to the surveys [19,12,1]).

Specifically, the rules that we start with are usually of the following type:

if  $x_1$  is  $A_1^j$  and  $x_2$  is  $A_2^j$  and... and  $x_n$  is  $A_n^j$ , then  $u$  is  $B^j$

where the  $x_i$  are parameters that characterize the plant,  $u$  is the control, and  $A_i^j, B^j$  are the natural language terms that are used to describe the  $j^{th}$  rule (e.g., “small”, “medium”, etc). The value  $u$  is a proper value for the control if and only if one of these rules is applicable; the property “ $u$  is a proper control” (which we will denote by  $C(u)$ ), therefore, can be described as follows:

$$\begin{aligned} C(u) \equiv & (A_1^1(x_1) \& A_2^1(x_2) \& \dots \& A_n^1(x_n) \& B^1(u)) \vee \\ & (A_1^2(x_1) \& A_2^2(x_2) \& \dots \& A_n^2(x_n) \& B^2(u)) \vee \\ & \dots \\ & (A_1^K(x_1) \& A_2^K(x_2) \& \dots \& A_n^K(x_n) \& B^K(u)) \end{aligned}$$

The natural language terms are described by membership functions, i.e., we describe  $A_i^j(x)$  as  $\mu_i^j(x)$ , the degree of belief that a given value  $x$  satisfies the property  $A_i^j$ ; similarly,  $B^j(u)$  is represented as  $\mu_j(u)$ . The logical connectives  $\&$  and  $\vee$  are interpreted in this context as the operations  $f_{\&}$  and  $f_{\vee}$  on degrees of belief. After these interpretations, we can form the membership function for control:  $\mu_C(u) = f_{\vee}(p_1, \dots, p_K)$ , where

$$p_j = f_{\&}(\mu_{j,1}(x_1), \mu_{j,2}(x_2), \dots, \mu_{j,n}(x_n), \mu_j(u)), \quad j = 1, \dots, K.$$

The system must supply a control, so we must end up with a single value  $u$  of the control that will actually be applied. An operation that transforms a membership function into a single value is called a *defuzzification*. To complete the fuzzy control methodology, therefore, we must apply some defuzzification operator  $F$  to the membership function  $\mu_C(u)$  and thus obtain the desired value  $\bar{u} = f_C(\vec{x})$  of the control that corresponds to  $\vec{x} = (x_1, \dots, x_n)$ .

We define these notions more carefully in the next section.

**How good can the result of fuzzy control be?** Of course, if the rules are bad, we end up with a bad control. By “bad” we mean that later on, when we learn how the plant works and compute the optimal control value, it will turn out that the control supplied by our fuzzy controller is much worse than this optimal one. The question we address is whether it is at all possible to have chosen *good rules*, i.e., does there exist (in principle) a set of rules that will lead us exactly to the optimal (or almost optimal) control? We answer this question in the affirmative in this paper, and we prove the correctness of our answer.

Of course, if we do not yet know the exact dynamics of the plant, we cannot tell which rules are good and which are bad, but at least, we hope that if we do choose the rules

correctly, we will get a very good control. In this sense, fuzzy control is a *universal control tool*, because an arbitrary control strategy (in particular, a control strategy that is in some sense optimal) can be obtained in principle by applying a fuzzy control methodology to some set of rules.

*Comment.* The fact that fuzzy control is a *universal control tool* does not necessarily mean that it is the *best* control tool. Indeed, many different tools, such as neural networks, polynomial functions etc, are also universal. Some of these tools may turn out to be better: e.g., they may require fewer computation time to compute the control. Universality of a control tool is, therefore, not *sufficient* for it to be a good tool, but universality is definitely a *necessary* requirement if we want a tool applicable to all possible control situations.

**What was already known.** The result that fuzzy control is a universal control tool has been proved earlier (e.g., [2,7,20,21]) for the case when the state of a plant can be described by finitely many parameters and a special type of fuzzy control methodology is used (in the next section we describe these results in some detail). In mathematical terms, this means that fuzzy controls can approximate an arbitrary real continuous function  $f(\vec{x})$  with arbitrary accuracy. This means that for a plant whose state can be described by finitely many parameters, for an arbitrary potential control strategy  $u(\vec{x})$ , and for an arbitrary accuracy  $\varepsilon > 0$ , a fuzzy controller can be designed to achieve the given control to the given accuracy.

**What these results do not cover.** These results do not cover many important control problems:

- First, they do not cover the control of distributed systems (i.e., plants that require infinitely many parameters to describe their current state). Can we apply fuzzy control to such systems?
- Second, the above results show that any given control of a given plant can be implemented as a fuzzy controller. The goals of control theory, however, are usually more ambitious than controlling one plant: the main goal is to design universal control methods, so that we can input the description of a plant and the desired objective and generate an appropriate control. The above-mentioned results show only that for each plant and for an arbitrary objective, optimal control can be obtained by a fuzzy control methodology (if we use appropriate rules). Can we design a *universal fuzzy controller*, i.e., can we formulate the rules that would enable us to transform the description of a plant and an objective into an actual control strategy?

**What we are planning to do.** In this paper, we prove that for an arbitrary fuzzy control methodology, fuzzy control is a universal tool for any plant (including plants that are *distributed systems*, i.e., plants whose state requires infinitely many parameters to describe). We also prove that there exists a *universal fuzzy controller*, i.e., one that inputs a description of a plant and the desired objective and generates an appropriate control. Mathematically, we prove that fuzzy systems can approximate arbitrary continuous functionals (some of our results first appeared in [15]).

**The structure of the paper.** In Section 2, we present the main definitions and results for plants with finitely many parameters. In Section 3, we explain the importance of these results for control theory, namely, that they show that an arbitrary potential (thus “realistic”) control strategy can be approximated by a fuzzy controller when the rules are chosen appropriately. In Section 4, we describe the control problems that are not directly covered by these results. In Sections 5 through 7, we show that fuzzy control is a universal tool for these problems as well. Section 8 contains the proofs.

## 2. DEFINITIONS AND THE MAIN RESULTS FOR PLANTS WITH FINITELY MANY PARAMETERS

### 2.1. General definition of a fuzzy control

#### Definition 1.

1. By a *fuzzy set*, we mean a function  $\mu : R \rightarrow [0, 1]$ . We say that a set is *non-empty* if this function is not identically 0 for all  $x \in R$ . This function will also be called a *membership function*.
2. A membership function that is equal to 1 for  $x = a$  and 0 for all other  $x$  is called *crisp*, or a (*crisp*) *number*. A crisp number will be denoted by  $\delta(x - a)$ .
3. A function  $\mu(x) = \exp(-(x - a)^2/k)$  is called a *Gaussian* membership function. In particular, the function  $\mu(x) = \exp(-x^2)$  is called a *basic Gaussian* membership function.
4. We assume that a finite set  $\mathcal{W}$  is given whose elements will be called *natural language words*. Elements of the set  $\mathcal{W}$  will be denoted by capital letters. We assume that a mapping is given that assigns to every  $W \in \mathcal{W}$  a fuzzy set  $\mu_W$ .

**Definition 2.** Assume that a positive integer  $n$  is given; this  $n$  will be called the *number of variables*.

1. By a *rule* we mean a formula of the type  $A_1(x_1) \& \dots \& A_n(x_n) \rightarrow B(u)$ , where  $A_i$  and  $B$  are natural language words (i.e., elements of  $\mathcal{W}$ ).
2. By a *rule base* (or *knowledge base*) we mean a finite set of rules. The number of rules in a rule base will be denoted by  $K$ , and the  $j^{th}$  rule will be denoted by  $A_{j1}(x_1) \& \dots \& A_{jn}(x_n) \rightarrow B_j(u)$ . A membership function that corresponds to  $A_{ji}$ , will be denoted by  $\mu_{ji}$ , and a membership function that corresponds to  $B_j$ , will be denoted by  $\mu_j$ .

#### Definition 3.

1. By an *&-operation* we mean a continuous function  $f_{\&} : [0, 1] \times [0, 1] \rightarrow [0, 1]$  that satisfies the following four properties:

- $f_{\&}(0, 0) = f_{\&}(0, 1) = f_{\&}(1, 0) = 0, f_{\&}(1, 1) = 1;$
- $f_{\&}(a, b) = f_{\&}(b, a)$  for all  $a, b;$
- $f_{\&}(a, b) \leq a$  for all  $a$  and  $b;$
- if  $a > 0$  and  $b > 0$ , then  $f_{\&}(a, b) > 0.$

2. By an  *$\vee$ -operation* we mean a continuous function  $f_{\vee} : [0, 1] \times [0, 1] \rightarrow [0, 1]$  that satisfies the following three properties:

- $f_{\vee}(0, 0) = 0, f_{\vee}(0, 1) = f_{\vee}(1, 0) = f_{\vee}(1, 1) = 1;$
- $f_{\vee}(a, b) = f_{\vee}(b, a)$  for all  $a, b;$

- $f(a, b) \geq a$  for all  $a$  and  $b$ .
3. By a *defuzzification procedure*  $F$  we mean a mapping that transforms a membership function  $\mu(x)$  into a number and satisfies the following properties:
- if  $\mu(x) = 0$  for all  $x \in (-\infty, a)$ , then  $F(\mu) \geq a$ ;
  - if  $\mu(x) = 0$  for all  $x \in (-\infty, a]$ , then  $F(\mu) > a$ ;
  - if  $\mu(x) = 0$  for all  $x \in (a, \infty)$ , then  $F(\mu) \leq a$ ;
  - if  $\mu(x) = 0$  for all  $x \in [a, \infty)$ , then  $F(\mu) < a$ .
4. By a *centroid defuzzification*  $F_c$  we mean a mapping that transforms  $\mu$  into

$$F(\mu) = \frac{\int u\mu(u) du}{\int \mu(u) du}.$$

*Comments.*

1. The above binary operations are slightly more general than the usual definitions of  $t$ -norms and  $t$ -conorms in the literature [5].
2. It is easy to check that a centroid defuzzification is an example of a defuzzification in the sense of this definition.
3. If a membership function is defined only at finitely many points, then integrals degenerate into sums. Thus, if we have a membership function  $\mu(x)$  that is different from 0 only for the values  $x_1, \dots, x_n$ , then

$$F_c(\mu) = \frac{x_1\mu(x_1) + \dots + x_n\mu(x_n)}{\mu(x_1) + \dots + \mu(x_n)}.$$

**Definition 4.** Assume that  $B$  is a rule base,  $f_{\&}$ ,  $f_{\vee}$  are  $\&$ - and  $\vee$ -operations, and  $F$  is a defuzzification procedure.

1. By a *fuzzy control*, we mean a function that maps every vector  $\vec{x} = (x_1, \dots, x_n) \in R^n$  into  $\bar{u} = F(\mu)$ , where  $\mu(u) = f_{\vee}(p_1(\vec{x}, u), p_2(\vec{x}, u), \dots, p_K(\vec{x}, u))$ ,  $p_j = f_{\&}[r_j(\vec{x}), \mu_j(u)]$ , and  $r_j(\vec{x}) = f_{\&}[\mu_{j1}(x_1), \dots, \mu_{jn}(x_n)]$ .

*Comments.*

1.  $r_j(u, \vec{x})$  is the degree of belief that rule  $j$  is applicable for the given values  $\vec{x}$  ( $r$  stands for “rule”).
2. In particular, if all  $B_j$  are crisp numbers (i.e., if  $\mu_j(u) = \delta(u - \bar{u}_j)$  for some  $\bar{u}_j$ ), all  $\bar{u}_j$  are different, and  $F = F_c$ , then

$$\bar{u} = \frac{\bar{u}_1 r_1(\vec{x}) + \dots + \bar{u}_K r_K(\vec{x})}{r_1(\vec{x}) + \dots + r_K(\vec{x})}.$$

This formula is much easier to compute than the formula with integrals; for this reason, it is sometimes used as a fuzzy control methodology even when the words  $B_j$  are not crisp numbers (i.e., we choose the values  $\bar{u}_j$  for which  $\mu_j(u)$  is the largest possible and act as if  $\mu_j = \delta(u - \bar{u}_j)$  (see, e.g., [12])).

3. It should be noted that, unlike the general result, the simplified formula from 2. does not depend on the choice of an  $\vee$ -operation.

## 2.2. Specific fuzzy control methodologies that are known to be universal tools

**Definition 5.** Fix a pairs of membership functions  $\mu_x$  and  $\mu_y$  and call them *basic membership functions*. A membership function  $\mu(x)$  is of type  $\mu_x(x)$  if  $\mu(x) = \mu_x(kx+l)$ , for some real numbers  $k \neq 0$  and  $l$ . A membership function  $\mu(x)$  is of type  $\mu_y(x)$  if  $\mu(x) = \mu_y(kx+l)$ , for some real numbers  $k \neq 0$  and  $l$ .

### Definition 6.

1. By a *fuzzy methodology*  $\mathcal{M}$  we mean a tuple consisting of basic membership functions  $\mu_x$  and  $\mu_y$ ,  $\&$ - and  $\vee$ -operations  $f_{\&}(a, b)$  and  $f_{\vee}(a, b)$ , and a defuzzification procedure  $F$ .
2. Assume that a fuzzy methodology  $\mathcal{M}$  is fixed. By  $\mathcal{F}(\mathcal{M})$  we denote the class of all control strategies that can be obtained using  $\mathcal{M}$ , i.e., the class of all functions  $f : R^n \rightarrow R$  that are equal to  $f(\vec{x}) = F(\mu_C)$ , where  $\mu_C(u) = f_{\vee}(p_1, p_2, \dots, p_j, \dots, p_K)$ ,  $p_j$  is as defined previously, all membership functions  $\mu_{j,i}(x)$  are of type  $\mu_x$ , and all membership functions  $\mu_j(u)$  are of type  $\mu_y$ .

**Denotations.** Assume that  $U \subset R^n$  is a domain in  $R^n$ .

1. By  $C(U)$  we denote the set of all continuous functions from  $U$  to  $R$ .
2. For every  $f \in C(U)$  we define its *sup-norm*  $\|f\|$  as

$$\|f\| = \sup_{\vec{x} \in U} |f(\vec{x})|.$$

3. For every function  $f : R^n \rightarrow R$ ,  $f|_U$  will be used to mean a restriction of  $f$  to the set  $U$ .
4. If  $\mathcal{F}$  is a class of functions, then by  $\mathcal{F}|_U$  we mean the set of restrictions  $f|_U$  for all functions  $f \in \mathcal{F}$ .

### Definition 7.

1. We say that a class of functions  $\mathcal{F}$  is *dense* in the sup-norm in  $C(U)$  if for every continuous function  $f \in C(U)$ , and for every  $\varepsilon > 0$ , there exists a function  $\tilde{f} \in \mathcal{F}$  such that  $\|f - \tilde{f}\| \leq \varepsilon$ .
2. We say that a fuzzy methodology  $\mathcal{M}$  is a *universal control tool for plants that are described by finitely many parameters* (or, for short, a *universal control tool*) if for every compact  $U \subset R^n$ , the corresponding class  $\mathcal{F}(\mathcal{M})|_U$  is dense in the sup-norm in  $C(U)$ .

### Comments.

1. If a set  $\mathcal{F}(\mathcal{M})$  of all the functions that can be obtained by using a certain methodology is dense in  $C(U)$ , then an arbitrary potential control strategy can be approximated by a function from  $\mathcal{F}$ . The statement really means, therefore, that this particular methodology is a universal control tool.

2. For the time being, we have omitted the words “for plants that are described by finitely many parameters” just to make all our formulations shorter. We will see later on,

however, that this omission makes perfect sense; specifically, we will prove that if a fuzzy methodology is a universal control tool for “finite” plants, then it is also a universal control tool for other plants as well.

3. Historically, the first theorem showing that fuzzy control is a universal control tool was proved in [21]:

**THEOREM [21].** *The methodology  $\mathcal{M}$  for which  $f_{\&}(a, b) = ab$ ,  $\mu_x$  is an everywhere positive function,  $\mu_y$  is crisp, and  $F$  is a centroid, is a universal control tool.*

*Comments.*

1. In other words, fuzzy systems with product inference, centroid defuzzification, and everywhere positive membership functions are capable of approximating any potential control function on a compact set to arbitrary accuracy.

2. One of the reasons why it is reasonable to choose  $ab$  for the  $\&$ -operation is that, according to the psychological literature [17,23], the functions  $\min(a, b)$  and  $ab$  provide the best known description of what we humans actually understand the word “and” to mean.

3. As a particular case, we get a similar result for Gaussian membership functions:

**THEOREM [20].** *A methodology  $\mathcal{M}$  for which  $f_{\&}(a, b) = ab$ ,  $\mu_x$  is Gaussian,  $\mu_y$  is crisp, and  $F$  is a centroid, is a universal control tool.*

*Comments.*

1. In some situations, Gaussian membership functions are indeed the most adequate ones (see, e.g., [9,10]).

2. A similar result (i.e., that using a certain fuzzy methodology makes fuzzy control a universal tool) was proved by Kosko [7].

3. These two theorems state that, for certain methodologies  $\mathcal{M}$ , fuzzy control is a universal control tool. We now describe our own results which show that this is also true if we use other methodologies.

### **2.3. A fuzzy methodology with Gaussian membership functions and with $\min$ as $\&$ is also a universal control tool**

We mentioned earlier that, according to the psychological data, two  $\&$ -operations are the most adequate:  $\min$  and  $ab$ . It has already been demonstrated that if we use the product, then fuzzy methodology becomes a universal tool. Let us now show that the same is true for  $\min$ .

**THEOREM 1.** *A methodology  $\mathcal{M}$  for which  $f_{\&}(a, b) = \min(a, b)$ ,  $\mu_x$  is Gaussian,  $\mu_y$  is crisp, and  $F$  is a centroid, is a universal control tool.*

*Comments.*

1. In this case, the class  $\mathcal{F}(\mathcal{M})$  consists of all functions  $f : R^n \rightarrow R$  of the form

$$f(\vec{x}) = \frac{\sum_{j=1}^K \bar{u}_j \min_{i=1, \dots, n} \mu_{j,i}(x_i)}{\sum_{j=1}^K \min_{i=1, \dots, n} \mu_{j,i}(x_i)},$$

where the  $\mu_{j,i}(x)$  are Gaussian functions, and the  $\bar{u}_j$  are arbitrary real numbers.

2. For the reader’s convenience, all the proofs are collected in the last section of this paper.

#### 2.4. Fuzzy control is a universal tool for an arbitrary realistic methodology

We now show that a similar result can be proven for an arbitrary choice of a defuzzification operation,  $\&$ – and  $\vee$ –operation, and basic membership function, as long as this choice is *realistic*. To explain this condition, we describe what is *not* exactly realistic about  $\mu_x$  being Gaussian and  $\mu_y$  being a (crisp) number:

- A Gaussian membership function is always positive. This means that if we use such a function to describe a natural language term like “small” or “medium”, even for arbitrary large values ( $10^6, 10^9$ , etc) there is a non-zero degree of belief that this particular value is small (or medium). Since this degree of belief is close to 0, it might be considered that this is nonetheless a good approximation to how we interpret words like “small”. If we want a really good description of the word’s natural language meaning, however, then we must take into consideration the fact that there are limit values beyond which no one would claim that this value is “small” (or, in general, that some value can be characterized by some particular word). Therefore, membership functions in general must be equal to 0 outside some interval.
- In the exposition above, we described  $\mu_y$  as a crisp membership function ( $\mu_j(u) = \delta(u - \bar{u}_j)$ ): this means that we know the exact control value  $u$ . In mathematical terms, if we make a slight change from  $\bar{u}_j$  to a nearby value  $u$ , our degree of belief “jumps” from 1 to 0. In real life, even if we know  $u$  pretty well, there is still some doubt about what exact control value to use. As a result, if we change the value  $u$  a little bit, the degree of belief should also change gradually. In other words, realistic membership functions must be *continuous*.

This leads us to the following definition:

**Definition 8.**

1. We say that a membership  $\mu$  function is *realistic* if it satisfies the following three conditions:

- $\mu$  is continuous;
- $\mu(x) > 0$  for all  $x$  in some interval  $(a, b)$ ;
- $\mu(x) = 0$  for all  $x$  outside  $(a, b)$ .

2. We say that a fuzzy methodology is *realistic* if all basic membership functions are realistic.

**THEOREM 2.** *An arbitrary realistic fuzzy methodology  $\mathcal{M}$  is a universal control tool.*

#### 2.5. Relationship between these results and similar results for neural networks

The results of this section are similar to previous results about the approximation capabilities of neural networks (see, e.g., [6,11]); this similarity is deeply-rooted mathematically (e.g., the mathematical techniques used in [20,21] for fuzzy control – the Stone-Weierstrass approximation theorem – is the same as was used in [6]). It is worth noting, however, that

in spite of the mathematical similarity, neural networks and fuzzy control are applicable to radically *different* real-life problems:

- If we have the *rules* formulated in natural language, then we can apply fuzzy control, but *not* neural networks.
- If, instead of rules, we have the recorded *experience* of successful controls, i.e., we know the values of the control applied in different situations, then we can use these *patterns* to teach a neural network. In this case, there is no way to apply fuzzy control (at least not directly).

Neural networks and fuzzy control are not competitors, then, but rather complementary techniques.

### 3. THESE UNIVERSALITY RESULTS ARE IMPORTANT FOR CONTROL

The theorems that we have just described are very important in applications of fuzzy control methodology to real-life control problems. Suppose we consider a plant whose state can be described by finitely many parameters  $(x_1, \dots, x_n)$  and whose possible controls can be characterized by  $p$  control parameters  $(u_1, \dots, u_p)$ . To design a control for such a system means to find a way to generate proper control values  $u_1, \dots, u_p$  for each state of the plant  $(x_1, \dots, x_n)$ . In mathematical terms, a control strategy is a function from  $R^n$  to  $R^p$ .

In real-life situations, there are a priori bounds on the value of every physical parameter: velocity cannot take a value that exceeds the speed of light, position cannot take values that exceed the size of the area under analysis, etc. If we denote by  $X_i$  the biggest possible value of  $|x_i|$ , then we must consider only the values  $x_i \in [-X_i, X_i]$ , and therefore, only the states  $(x_1, \dots, x_n)$  from a compact set

$$S = [-X_1, X_1] \times [-X_2, X_2] \times \dots \times [-X_n, X_n]$$

are physically possible. We need to consider, therefore, only functions from  $S$  to  $R^p$ .

In non-linear control theory (see, e.g., [14]), for many reasonable objectives, the optimal control is not continuous (the so-called “bang-bang” control is a typical example of this phenomenon). However, when it comes to practical implementations, we have to take into consideration the fact that every real-life device, no matter how fast it is, produces a *continuous* change of parameters: we cannot immediately change position, we cannot change velocity in zero time (with infinite acceleration), and even the change in electric current (that occurs, e.g., when we switch something on or off) is continuous. Therefore, any real-life (hardware) control is a *continuous* function from  $S$  to  $R^p$ .

If we take into consideration restriction related to hardware (in particular, continuity), then the problem of finding an optimal control for a given hardware device can be formulated as that of optimizing an objective function under the condition that a *continuous* function  $f : S \rightarrow R^p$  satisfies given (hardware-related) restrictions. As a result, if we take implementation restrictions into consideration, the “ideal” (i.e., optimal) control will be continuous.

Since we are talking about a hardware control, and hardware devices cannot be absolutely precise, we cannot guarantee that the actual control will precisely coincide with the function  $f : S \rightarrow R^p$  that is prescribed by (this restricted) optimization. However, by imposing more and more strict conditions on the quality of the hardware parts, we can guarantee that the resulting hardware control is as close to the theoretical one as possible.

Given these observations, Theorem 2 really asserts that for an arbitrary “ideal” control  $f : S \rightarrow R^p$  (not known to us), and for an arbitrary  $\varepsilon > 0$ , there exists a fuzzy controller that generates  $f$  with accuracy  $\varepsilon$ . Whatever ideal control may exist, in other words, there also exists a fuzzy controller that is as close to it as possible.

#### 4. THE PREVIOUS APPROXIMATION RESULTS DO NOT COVER ALL POSSIBLE CONTROL SITUATIONS

Theorems 1 and 2 cover many important control problems, but they do not cover all of them:

- Not all systems can be described by finitely many parameters. For example, a proper control decision for a chemical reaction requires knowledge of the temperature  $T$ , density  $\rho$ , and other characteristics at all the points  $\vec{u}$  inside the reactor. In this case, to describe a state of the plant, we must know the functions  $T(\vec{u})$ ,  $\rho(\vec{u})$ , etc. If we are controlling a plant that consist of several reactors, we need to know the functions  $T(\vec{u})$ ,  $\rho(\vec{u}), \dots$ , for  $\vec{u}$  for the first reactor  $U_1$ , and also similar functions to describe other reactors  $U_2, \dots$ . Such systems are called *systems with distributed parameters*, or *distributed systems*. Can we apply fuzzy control to such systems?
- Theorems 1 and 2 address control of a plant under one specific control objective. In real life, however, we may have different plants and different control objectives. Therefore, the objective of control theory is not only to provide optimal control of specific plants under specific criteria, but also to develop general methods that would help to control an arbitrary plant under an arbitrary objective. We would like fuzzy methodology to provide us with a *universal controller* in the following sense: we supply it with the description of the system and with an objective (i.e., an optimality criterion), and it will generate the optimal control for this very system under this very criterion.

#### 5. FUZZY SYSTEMS CAN APPROXIMATE CONTROLS OF DISTRIBUTED SYSTEMS: MAIN RESULT

##### 5.1. How to describe a state of a distributed system

**Definition 9.** Suppose that  $n$  and  $m$  are non-negative integers, and  $U_1, \dots, U_m$  are compact sets with metrics  $\rho_1, \dots, \rho_m$ . By a *state* we mean a tuple  $s = (z_1, \dots, z_n, f_1, \dots, f_m)$ , where  $z_i$  are real numbers, and  $f_j$  is a continuous function from  $U_j$  to  $R$ .

##### Denotations.

1. By  $C(U_j)$  we will denote the set of all continuous functions from  $U_j$  to  $R$ ; a topology on  $C(U_j)$  is determined by a  $C$ -metric  $\|f - g\| = \sup_{u \in U_j} |f(u) - g(u)|$ .

2. The set  $\mathcal{S}$  of all states will be denoted by  $R^n \times C(U_1) \times \dots \times C(U_m)$ . Let us define the following metric on  $\mathcal{S}$ :

$$\rho(s, \tilde{s}) = \max(|z_1 - \tilde{z}_1|, \dots, |z_n - \tilde{z}_n|, \|f_1 - \tilde{f}_1\|, \dots, \|f_m - \tilde{f}_m\|),$$

where  $s = (z_1, \dots, z_n, f_1, \dots, f_m)$  and  $\tilde{s} = (\tilde{z}_1, \dots, \tilde{z}_n, \tilde{f}_1, \dots, \tilde{f}_m)$ .

**Definition 10.** By a *set of all physically possible states of a plant with distributed parameters* (or, for short, a *set of all physically possible states*), we will mean a compact set  $S \subset \mathcal{S}$ . States from  $S$  will be called *physically possible*.

*Remark.* These definitions are different from the ones that a reader may encounter in the majority of the papers on distributed parameter control. The main reason for this difference is that we are trying to be realistic, in particular, we are trying to take into consideration hardware-related restrictions on the control strategies. In real life, as we have already mentioned, all physical quantities are bounded. The coordinates of a controlled object (*plant*) are bounded by the area we can control, its velocity is bounded by the velocity of light, etc. For an arbitrary physical quantity, then, its values are bounded, and the values of its derivatives are also bounded by some number  $M$ . Hence, the functions  $f_j(u)$  that correspond to physically possible states are uniformly bounded, and their derivatives are uniformly bounded.

Given the boundedness of realistic control, therefore, we can use the known description of compact sets in a space  $C(U)$  of continuous functions from a compact space  $U$  to  $R$ . According to the Arzelá-Ascoli Theorem (see, e.g., [18]), a set of functions  $\mathcal{F} \subseteq C(U)$  is pre-compact if and only if it is uniformly bounded (iff there exists a number  $M$  that bounds all of them, i.e., such that  $|f(u)| \leq M$  for all  $u \in U$  and for all  $f \in \mathcal{F}$ ) and uniformly continuous (i.e., for every  $\varepsilon$ , there exists a  $\delta$  such that if  $d(u, u') \leq \delta$  and  $f \in \mathcal{F}$ , then  $|f(u) - f(u')| \leq \varepsilon$ ). If a set is pre-compact, then its closure is compact.

We have argued that the functions that correspond to physical states are uniformly bounded, and that their derivatives are uniformly bounded. This means that they are uniformly continuous and, therefore, that the set of all the functions that correspond to physical states is precompact. If we add limits to this set we get a *compact* set of possible states.

## 5.2. How to describe control strategies, and how we can apply fuzzy control to distributed systems

**Definition 11.** Suppose that an integer  $p$  is given; elements  $u = (y_1, \dots, y_p)$  will be called *control values*. On  $R^p$ , we define a metric  $\|y - \tilde{y}\| = \max_l |y_l - \tilde{y}_l|$ . By a *control strategy* we mean a continuous function  $J : S \rightarrow R^p$  that assigns to every physically possible state  $s$  a control value  $J(s)$ .

*Comment.* We are planning to show that an optimal control strategy  $J$  can be (in principle) approximated as closely as possible by an appropriate fuzzy controller. To do this, we need to prove that fuzzy control methodology is a universal control tool for distributed systems as well. In order to formulate a relevant mathematical statement, we must do two things:

- first, we must describe what it means for two control strategies to be close;
- second, we must describe how to apply fuzzy control methodology to distributed systems.

The first problem is solved by the following natural definition:

**Definition 12.** Suppose that  $\varepsilon > 0$ . We say that control strategies  $J$  and  $\tilde{J}$  are  $\varepsilon$ -close if for all  $s \in S$ ,  $\|J(s) - \tilde{J}(s)\| \leq \varepsilon$ .

The remaining problem is how to apply fuzzy control methodology to distributed systems. To answer this question, let us analyze how a human operator chooses his control decision. Generally speaking, the operator's decision is based on his knowledge of the current state which can come either from measurements or from his expert estimate of certain characteristics of the state. For example, when driving a car, we use the speedometer to determine speed and rely on our "expert" ability to estimate the distance between our car and other cars. Even if an expert has a potentially unlimited number of measurement devices at his disposal, he still can use only finitely many measurement or estimation results. For instance, if he is controlling a chemical reactor, then he can only measure the temperature at finitely many points  $\vec{u}$ . As a result, he has only finitely many parameters on which to base his decisions. For such a case, we already know how to formulate the rules, and how to apply a fuzzy control methodology.

For the case of a chemical reactor, we can formulate the following rule: if all the sensors on the surface of the reactor show that the temperature is big (i.e., drastically exceeds the norm), then we must cut the influx of the reacting products to small. If we denote the number of temperature sensors by  $N$ , and their locations by  $\vec{u}^{(1)}, \vec{u}^{(2)}, \dots, \vec{u}^{(N)}$ , then this rule can be formulated as follows:

if  $T(\vec{u}^{(1)})$  is big,  $T(\vec{u}^{(2)})$  is big, ..., and  $T(\vec{u}^{(N)})$  is big, then  $u$  is small.

This rule is similar to the ones that we described when we talked about fuzzy control using the fuzzy properties "big" and "small", the only difference being that here we have the values  $x_1 = T(\vec{u}^{(1)})$ , ...,  $x_N = T(\vec{u}^{(N)})$  as input parameters for control. Let us describe this idea in general terms.

**Definition 13.** Assume that a set  $S \subset \mathcal{S} = R^n \times C(U_1) \times \dots \times C(U_m)$  of all physically possible states is given. By a *rule*, we mean a formula of the type  $A_1(x_1) \& \dots \& A_n(x_n) \rightarrow B(u)$ , where the  $A_i$  and  $B$  are natural language words (i.e., elements of  $\mathcal{W}$ ), and each value  $x_i$  is either  $z_i$ , or  $f_j(u_j)$  for some  $j$  and some  $u_j \in U_j$ . By a *rule base* (or *knowledge base*) we mean a finite set of rules.

*Comment.* Every rule base contains only finitely many rules, and every rule uses only the values of  $f_j$  at finitely many points  $u_j \in U_j$ . In other words, the rules describe how the control value  $u$  depends on finitely many parameters  $z_1, \dots, z_n, f_1(u_1^{(1)}), \dots, f_1(u_1^{(N_1)}), \dots, f_m(u_m^{(1)}), \dots, f_m(u_m^{(N_m)})$  for finitely many values  $u_1^{(1)} \in U_1, \dots, u_1^{(N_1)} \in U_1, \dots, u_m^{(1)} \in U_m, \dots, u_m^{(N_m)} \in U_m$ . Therefore, if we have already chosen a

fuzzy control methodology  $\mathcal{M}$ , we can apply this methodology to this set of rules and get a control strategy that is a function of these same parameters:

$$J_{\mathcal{M}}(s) = J_{\mathcal{M}}(z_1, \dots, z_n, f_1, \dots, f_m) = \\ J_{\mathcal{M}}(z_1, \dots, z_n, f_1(u_1^{(1)}), \dots, f_1(u_1^{(N_1)}), \dots, f_m(u_m^{(1)}), \dots, f_m(u_m^{(N_m)})).$$

The question is: suppose that the optimal control (unknown to us) is described by a function  $J(s)$  and that the accuracy with which we want to get this control is  $\varepsilon$ . If we restrict ourselves only to control strategies that stem from applying fuzzy control methodology, will it be possible for us to achieve this optimal control (using an appropriate set of rules), i.e., is the described fuzzy control methodology a universal tool for distributed systems as well?

**Definition 14.** Suppose that  $S$  is a set of physically possible states of a plant with distributed parameters,  $\mathcal{M}$  is a fuzzy control methodology, and  $\mathcal{R}$  is a rule base. The result of applying  $\mathcal{M}$  to  $\mathcal{R}$  is defined as in Definition 4.

*Comment.* In other words, the resulting control strategy is equal to  $J(s) = F(\mu_C)$ , where  $\mu(u) = f_{\vee}(p_1(\vec{x}, u), p_2(\vec{x}, u), \dots, p_K(\vec{x}, u))$ ,  $p_j = f_{\&}[r_j(\vec{x}), \mu_j(u)]$ ,  $r_j(\vec{x}) = f_{\&}[\mu_{j1}(x_1), \dots, \mu_{jn}(x_n)]$ , and each  $x_i$  is either equal to  $z_{i'}$ , or to  $u_j^{(i')}$  for some  $i'$  and  $j$ .

**Definition 15.** We say that a fuzzy control methodology  $\mathcal{M}$  is a *universal control tool for systems with distributed parameters* if for every set  $S$ , for every control strategy  $J : S \rightarrow R$ , and for every  $\varepsilon > 0$ , there exists a set of rules for which  $\mathcal{M}$  generates a control strategy  $\tilde{J}$  that is  $\varepsilon$ -close to  $J$ .

**THEOREM 3.** *If a fuzzy control methodology  $\mathcal{M}$  is a universal control tool for plants that are described by finitely many parameters, then this methodology is also a universal control tool for systems with distributed parameters.*

*Comments.*

1. From Theorem 3, it follows that all methodologies that were described in the previous section as universal tools for plants with finitely many parameters are also universal control tools for systems with distributed parameters.

2. In purely mathematical terms, this result can be reformulated as follows:

**Definition 16.** Assume that  $n, m$  and  $p$  are positive integers,  $U_1, \dots, U_m$  are compact sets,  $S \subset R^n \times C(U_1) \times \dots \times C(U_m)$  is a compact set, and  $\mathcal{M}$  is a fuzzy control methodology. By  $\mathcal{J}(\mathcal{M})$  we denote the set of all functionals  $J : S \rightarrow R$  that can be obtained by applying  $\mathcal{M}$  to different rule bases.

**THEOREM 3'.** *Assume that  $n, m$  and  $p$  are positive integers,  $U_1, \dots, U_m$  are compact sets,  $S \subset R^n \times C(U_1) \times \dots \times C(U_m)$  is a compact set,  $J : S \rightarrow R^p$  is a continuous functional, and  $\varepsilon > 0$  is a real number. Then, there exists a set of rules such that applying  $\mathcal{M}$  to it yields a functional  $\tilde{J} \in \mathcal{J}(\mathcal{M})$  that is  $\varepsilon$ -close to  $J$  on  $S$ .*

In other words, functionals generated by fuzzy control are dense in the set of all possible functionals.

## 6. TAKING LOCATION ERROR INTO CONSIDERATION

*Remark.* Theorem 3 corresponds to the case when we can measure the values of  $f_j$  precisely at given points  $u_j^{(i)}$ . In real life, we can only guarantee the locations  $u_j^{(i)}$  with some finite accuracy  $\varepsilon_u$ . We now prove that even if we take this inaccuracy into consideration, fuzzy control methodology will still be a universal control tool.

**Definition 17.** Suppose that  $\varepsilon > 0$  and  $\varepsilon_u > 0$ . We say that for a location error  $\varepsilon_u$ , a control strategy  $J$  is  $\varepsilon$ -approximable by a fuzzy controller if for every  $j$  from 1 to  $m$ , there exist elements  $u_j^{(1)}, \dots, u_j^{(N_j)} \in U_j$  and there exists a function  $\tilde{J} \in \mathcal{J}(\mathcal{M})$  such that for every  $s = (z_1, \dots, z_n, f_1, \dots, f_m) \in S$ , and for every  $\tilde{u}_j^{(i)}$  such that  $\rho_j(u_j^{(i)}, \tilde{u}_j^{(i)}) \leq \varepsilon_u$ , we have

$$\|J(s) - \tilde{J}(z_1, \dots, z_n, f_1(\tilde{u}_1^{(1)}), \dots, f_1(\tilde{u}_1^{(N_1)}), \dots, f_m(\tilde{u}_m^{(1)}), \dots, f_m(\tilde{u}_m^{(N_m)}))\| \leq \varepsilon.$$

**THEOREM 4.** For an arbitrary control strategy  $J$  and for an arbitrary  $\varepsilon > 0$  there exists a positive real number  $\varepsilon_u$  such that for location error  $\varepsilon_u$ ,  $J$  is  $\varepsilon$ -approximable by a fuzzy controller.

This definition and theorem can also be reformulated in purely mathematical terms, viz.

**Definition 16'.** Assume that  $n, m$  and  $p$  are positive integers,  $S \subset R^n \times C(U_1) \times \dots \times C(U_m)$  and  $J : S \rightarrow R^p$  is a mapping. Suppose also that  $\varepsilon > 0$  and  $\varepsilon_u > 0$  are real numbers (the value  $\varepsilon_u$  will be called a *location error*). We say that for location error  $\varepsilon_u$ , a mapping  $J$  is  $\varepsilon$ -approximable by a fuzzy system if for every  $j$  from 1 to  $m$ , there exist values  $u_j^{(1)}, \dots, u_j^{(N_j)} \in U_j$  and there exists a function  $\tilde{J} \in \mathcal{J}(\mathcal{M})$  such that for every  $(z_1, \dots, z_n, f_1, \dots, f_m) \in S$ , and for every  $\tilde{u}_j^{(i)}$  such that  $\rho_j(u_j^{(i)}, \tilde{u}_j^{(i)}) \leq \varepsilon_u$ , we have

$$\|J(s) - \tilde{J}(z_1, \dots, z_n, f_1(\tilde{u}_1^{(1)}), \dots, f_1(\tilde{u}_1^{(N_1)}), \dots, f_m(\tilde{u}_m^{(1)}), \dots, f_m(\tilde{u}_m^{(N_m)}))\| \leq \varepsilon.$$

**THEOREM 4'.** Assume that  $n, m$  and  $p$  are positive integers,  $S \subset R^n \times C(U_1) \times \dots \times C(U_m)$  is a compact set,  $J : S \rightarrow R^p$  is a continuous functional, and  $\varepsilon > 0$  is a real number. Then there exists a real number  $\varepsilon_u > 0$  such that for location error  $\varepsilon_u$ ,  $J$  is  $\varepsilon$ -approximable by a fuzzy system.

## 7. THE EXISTENCE OF A UNIVERSAL FUZZY CONTROLLER

*Remark.* By a universal control strategy we mean a method that, given:

- a description of a plant (i.e., a description of how its states are changed if we apply this or that control);
- an objective (i.e., a function of a trajectory that we are trying to minimize);
- a current state;

describes what control to apply. We already know how to describe a state, so let us show how to describe the plant's dynamics and our objectives.

Our definitions are slightly different from the ones that a reader may encounter in the majority of the papers on control theory. The main difference is that (as in the case of distributed control) we are trying to be realistic; in particular, we are trying to take into consideration hardware-related restrictions on the control strategies. Therefore, we restrict ourselves to *continuous* strategies only, and we assume that the set of all physically possible control strategies is compact. Specific explanations for the differences between our definitions and the usual ones will be provided after each definition.

### 7.1. How to describe the dynamics

**Definition 17.** By a *set of physically possible control values* we mean a compact set  $C \subset R^p$ . Control values from the set  $C$  will be called *physically possible*.

*Remark.* The motivation for this definition is the same as for Definition 10.

**Definition 18.** Suppose that  $S \subset R^n \times C(U_1) \times \dots \times C(U_m)$  is the set of physically possible states, and  $C$  is the set of physically possible control values. By a *dynamics* we mean a tuple  $d = (g_1, \dots, g_n, h_1, \dots, h_m)$  of continuous functions, where  $g_i : S \times C \rightarrow R$ , and  $h_j : U_j \times S \times C \rightarrow R$ .

*Remark.* A dynamics describes how a state evolves in time: if we are in the state  $s = (z_1, \dots, z_n, f_1, \dots, f_m) \in S$ , and apply control  $y \in C$ , then  $dz_i/dt = g_i(s, y)$  and  $df_j(u_j)/dt = h_j(u_j, s, y)$ .

**Definition 19.** On the set  $\mathcal{D} = C(S \times C)^n \times C(U_1 \times S \times C) \times \dots \times C(U_m \times S \times C)$  of all dynamics we can define a metric as follows: if  $d = (g_1, \dots, g_n, h_1, \dots, h_m)$ , and  $\tilde{d} = (\tilde{g}_1, \dots, \tilde{g}_n, \tilde{h}_1, \dots, \tilde{h}_m)$ , then

$$\rho(d, \tilde{d}) = \max(\max_i \|g_i - \tilde{g}_i\|, \max_j \|h_j - \tilde{h}_j\|),$$

where  $\|\cdot\|$  denotes a sup-norm.

**Definition 20.** By a *set of physically possible dynamics* we mean a compact set  $D \subset \mathcal{D}$ . Elements of  $D$  will be called *physically possible dynamics*.

*Remark.* The reason for this definition is the same as above: in real life situations, the rate with which we can change the state is limited ( $|g_i| \leq M$ ) and the difference between the rates at nearby points is also limited. The set of all physically possible dynamics is, therefore, uniformly bounded, equicontinuous and (hence) compact.

### 7.2. How to describe trajectories and objectives

**Definition 21.** Suppose that  $S \subset R^n \times C(U_1) \times \dots \times C(U_m)$  is the set of physically possible states, and  $T_0 > 0$  is a real number. By a *trajectory* we mean a tuple  $d = (G_1, \dots, G_n, H_1, \dots, H_m)$  of continuous functions, where

$$G_i : [0, T_0] \rightarrow R, \text{ and } H_j : [0, T_0] \times U_j \rightarrow R.$$

*Remark.* These functions are used to provide a temporal component to the state description; thus, at an arbitrary moment of time  $t \in [0, T_0]$ , the state  $s(t)$  is equal to  $(G_1(t), \dots, G_n(t), H_1(u_1, t), \dots, H_m(u_m, t))$ .

**Definition 22.** On the set  $\mathcal{T} = C([0, T_0])^n \times C([0, T_0] \times U_1) \times \dots \times C([0, T_0] \times U_m)$  of all trajectories we can define a metric as follows: if  $\tau = (G_1, \dots, G_n, H_1, \dots, H_m)$ , and  $\tilde{\tau} = (\tilde{G}_1, \dots, \tilde{G}_n, \tilde{H}_1, \dots, \tilde{H}_m)$ , then

$$\rho(\tau, \tilde{\tau}) = \max(\max_i \|G_i - \tilde{G}_i\|, \max_j \|H_j - \tilde{H}_j\|),$$

where  $\|\cdot\|$  denotes a sup-norm.

**Definition 23.** By a *set of physically possible trajectories* we mean a compact set  $T \subset \mathcal{T}$ . Elements of  $T$  will be called *physically possible trajectories*.

*Remark.* Physically possible trajectories are limited relative to what the states can be and relative to the rate at which the states change; the set of all physically possible trajectories, therefore, is compact.

**Definition 24.** By an *objective* we mean a continuous mapping  $o : T \rightarrow R$ .

*Remark.* In other words, an objective is a function that describes some quality of each physically possible trajectory. This can be the time that was necessary to reach the goal, the total fuel spent, etc. The goal of the control theory is to maintain the trajectory with the smallest possible value of an objective.

**Definition 25.** On the set  $\mathcal{O} = C(T)$  of all objectives we can define a  $C$ -metric.

**Definition 26.** By a *set of physically meaningful objectives* we mean a compact set  $O \subset \mathcal{O}$ . Elements of  $O$  will be called *physically meaningful objectives*.

### 7.3. What a universal control strategy is and how it can be represented by a fuzzy controller

**Definition 27.** By a *universal control strategy* we mean a continuous mapping

$$u : D \times O \times S \rightarrow R^p.$$

*Remarks.*

1. In other words, a universal control strategy takes a (description of the) dynamics, an objective, and a current state, and generates control values that take into consideration the plant being controlled and the objective. Every general method of control theory is an example of a universal control strategy (e.g., linearization methods, etc).

2. The notion of a universal control strategy is a reasonable theoretical concept that formalizes the main objective of control theory. To avoid misunderstanding, however, we

must emphasize that currently, for most interesting problems, no methods that implement universal strategies are known. In other words, for these problems, there is no *analytical* (or *algorithmic*) methodology that would help us find an optimal control based on given descriptions of a plant and of an objective. It is desirable, therefore, to *design* such a universal strategy by using the experience of experts who have been designing optimal controllers. In other words, we would like to formulate their experience in terms of rules, apply fuzzy control methodology to these rules, and derive a *universal controller*.

From this point of view, fuzzy control methodology is applied, as it were, on a *metalevel*: we are *not* using it to get a control value for a specific plant but rather to generate a universal algorithm that would enable us to generate control values for different plants.

Before we start this process, it would be nice to know whether such a universal fuzzy controller is at all possible, i.e., whether there exists a rule base that if we apply a fuzzy control methodology to it we will obtain a universal controller that is arbitrarily close to the (unknown) optimal one. Since we do not know the optimal universal control strategy, the only way to guarantee that such a rule base exists for the strategy we are interested in is to prove that such a rule base exists for an *arbitrary* universal control strategy. This is exactly what we propose to prove in this section.

**Definition 28.** Suppose that  $\varepsilon > 0$ . We say that universal control strategies  $u$  and  $\tilde{u}$  are  $\varepsilon$ -close if for all  $d \in D$ ,  $o \in O$ , and  $s \in S$ ,  $\|u(d, o, s) - \tilde{u}(d, o, s)\| \leq \varepsilon$ .

*Remark.* Suppose that we have an expert who knows how to design an optimal (or close to optimal) controller. In some situations this expert may just be using known formulas, while in other situations, he may be using his intuition and his experience. How can we extract this intuition from him?

We approach this problem via the following argument (which is similar to one for distributed systems in Section 5.2): Suppose that this super-expert is presented with the general description of a plant (i.e., its dynamics  $d$ ), an objective  $o$ , and a current state  $s = (z_1, \dots, z_n, f_1, \dots, f_m)$ , and he generates from them an optimal (or close to optimal) control decision  $u$ . In the course of generating this decision, the expert can use only finitely many values: he can use the results of applying dynamics  $d$  to some states; he can use the results of applying the objective to certain trajectories, and he can also measure the current state  $f_j(u_j)$  for some  $u_j \in U_j$ . Whatever rules he can formulate, therefore, will be formulated in terms of finitely many numbers of these types.

**Definition 29.**

1. Assume that we are given sets  $D$ ,  $O$ , and  $S$ . By a *rule for the universal fuzzy controller*, we mean a formula of the type  $A_1(x_1) \& \dots \& A_n(x_n) \rightarrow B(u)$ , where  $A_i$  and  $B$  are words from natural language (i.e., elements of  $\mathcal{W}$ ), and each value  $x_i$  is one of the following:

- $z_i$  for some  $i$ ;
- $f_j(u_j)$  for some  $j$  and some  $u_j \in U_j$ ;
- $g_i(s, c)$  for some state  $s \in S$  and control value  $c \in C$ ;
- $h_j(u_j, s, c)$  for some  $u_j \in U_j$ ,  $s \in S$ , and  $c \in C$ ;
- $o(\tau)$  for some trajectory  $\tau \in T$ .

2. By a *rule base* (or *knowledge base*), we mean a finite set of rules.

*Comment.* If we have a rule base  $\mathcal{R}$  and a fuzzy control methodology  $\mathcal{M}$ , then we can apply  $\mathcal{M}$  to  $\mathcal{R}$  and end up with a function  $F(x_1, \dots, x_n)$ . If we substitute the values  $x_i$  into this function we arrive at a universal control strategy

$$\tilde{u}(d, o, s) = F(\{g_i(s^{(k)}, c^{(k)})\}, \{h_j(u_j^{(k)}, s^{(k)}, c^{(k)})\}, \{o(\tau^{(k)})\}, \{x_i\}, \{f_j(u_j^{(k)})\})$$

for some states  $s^{(k)}$ , control values  $c^{(k)} \in C$ , elements  $u_j^{(k)} \in U_j$ , and trajectories  $\tau^{(k)}$ . Can an arbitrary universal control strategy be represented this way? The answer (in the affirmative) is given by the following theorem:

#### 7.4. Result: there exists a universal fuzzy controller

**THEOREM 5.** *Assume that a fuzzy methodology  $\mathcal{M}$  is a universal control tool,  $u$  is a universal control strategy, and  $\varepsilon > 0$ . Then there exists a rule base  $\mathcal{R}$  such that if we apply  $\mathcal{M}$  to it the resulting fuzzy control  $\tilde{u}$  will be  $\varepsilon$ -close to  $u$ .*

*Remark.* This result means that whatever general method of solving control problems someone invents, there is always the possibility of implementing this method by applying fuzzy control methodology to an appropriate set of rules. The resulting controller is *universal* in the sense that it will work for an arbitrary plant and an arbitrary objective.

## 8. PROOFS

**Proof of Theorem 1.** As in [20], it suffices to verify the hypotheses of the Stone-Weierstrass theorem.

The fact that  $\mathcal{F}(\mathcal{M})|_U$  is an algebra of functions follows simply from the fact that if the  $a_i$  and  $b_j$  are all positive, then

$$\min_{i=1, \dots, p} a_i \times \min_{j=1, \dots, q} b_j = \min_i \min_j a_i b_j.$$

Indeed, if  $f(\vec{x}) = [\sum_{j=1}^K \bar{u}_j \min(\mu_{j,i}(x_i), i = 1, \dots, n)] / [\sum_{j=1}^K \min(\mu_{j,i}(x_i), i = 1, \dots, n)]$  and  $g(\vec{x}) = [\sum_{j=1}^L \bar{v}_j \min(\nu_{j,i}(x_i), i = 1, \dots, n)] / [\sum_{j=1}^L \min(\nu_{j,i}(x_i), i = 1, \dots, n)]$ , and  $\alpha$  is a real number, then it is obvious that  $\alpha f(\vec{x})$  is in  $\mathcal{F}(\mathcal{M})|_U$ . The sum  $f(\vec{x}) + g(\vec{x})$  and product  $f(\vec{x})g(\vec{x})$  of the functions  $f$  and  $g$  involve terms of the form

$$\min(\mu_{j,i}(x_i), i = 1, \dots, n) \min(\nu_{k,i}(x_i), i = 1, \dots, n).$$

The last term is equal to  $\min(\mu_{j,i}(x_i)\nu_{k,i}(x_i), i = 1, \dots, n)$ , because the  $\mu_{j,i}(x_i)$ 's and the  $\nu_{k,i}(x_i)$ 's are positive. The result follows from the obvious fact that products of Gaussians are Gaussians.

$\mathcal{F}(\mathcal{M})|_U$  vanishes at no point:  $\min(a_i, i = 1, \dots, p)$  is always positive.

$\mathcal{F}(\mathcal{M})|_U$  separates points: for  $\vec{y}, \vec{z} \in R^n$ , consider the element of  $\mathcal{F}(\mathcal{M})|_U$ :  $f(\vec{x}) =$

$[\min_i(\exp(-1/2(x_i - y_i)^2))]/[\min_i(\exp(-1/2(x_i - y_i)^2) + \min_i(\exp(-1/2(x_i - z_i)^2))]$ .  
Then  $f(\vec{z})/f(\vec{y}) = \min_i(\exp(-1/2(z_i - y_i)^2)) = 1$  if and only if  $z_i = y_i$  for all  $i$ , which is impossible if  $\vec{y} \neq \vec{z}$ . Q.E.D.

## Proof of Theorem 2.

*Comment.* This proof will *not* use the Stone-Weierstrass Theorem, because it requires that the set of functions  $\mathcal{F}(\mathcal{M})$  be an *algebra*, i.e., that it be closed under addition and multiplication. For Gaussian functions, this is true, but for arbitrary membership functions it is not. We cannot use this ready-made result, therefore, and so we must invent a new proof.

Since  $U$  is compact, there exist a finite  $(\delta/2)$ -net, i.e., a finite set of points  $\vec{x}^1, \dots, \vec{x}^K \in U$  such that for every  $\vec{x} \in U$  there exists a  $j$  for which  $\rho(\vec{x}, \vec{x}^j) \leq \delta/2$ . Let us fix such a net.

Let  $g$  belong to  $C(U)$ , and  $\varepsilon > 0$ . Let us form a rule base with  $K$  rules, one rule for each point  $\vec{x}^j = (x_1^j, x_2^j, \dots, x_n^j)$  from the chosen  $\delta/2$ -net. Each rule will take the form (for point  $\vec{x}^j$ )

$$A_1^j(x_1) \& A_2^j(x_2) \& \dots A_n^j(x_n) \rightarrow B^j(u),$$

where the corresponding membership functions are  $\mu_{j,i}(x) = \tilde{\mu}_x((x - x_i^j)/\delta)$  and  $\mu_j(u) = \tilde{\mu}_y((u - u^j)/(\varepsilon/2))$ , where  $u^j = g(\vec{x}^j)$ , where  $\tilde{\mu}_x(x) = \mu_x(x(b - a)/2 + (a + b)/2)$  is a function of type  $\mu_x$  that is  $> 0$  for  $x \in [-1, 1]$ , and where  $\tilde{\mu}_y$  is defined similarly from  $\mu_y$ .

It suffices to show that  $\mu_C(u)$  is not identically 0 and that it is zero outside  $[g(\vec{x}) - \varepsilon, g(\vec{x}) + \varepsilon]$ ; the resulting inequality will then follow from the definition of the defuzzification procedure  $F$ .

First, we prove that  $\mu_C(u)$  is not identically 0. Since we chose the set  $\{\vec{x}^j\}$  as a  $\delta/2$ -net, there exists a  $j$  such that  $|x_i^j - x_i| \leq \delta/2$  for all  $i$ . Therefore,  $|x_i^j - x_i| < \delta$  which means that  $x_i \in (x_i^j - \delta, x_i^j + \delta)$ , and so  $\mu_{j,i}(x_i) > 0$  for all  $i$ . If we take  $u = g(\vec{x}^j)$ , we conclude that  $\mu_j(u) > 0$ . Since we required that  $f_{\&}(a, b) > 0$  if  $a > 0$  and  $b > 0$ , we can conclude that

$$p_j = f_{\&}(\mu_{j,1}(x_1), \mu_{j,2}(x_2), \dots, \mu_{j,n}(x_n), \mu_j(u)) > 0$$

Since  $f_{\vee}(a, b) \geq \max(a, b)$ , we conclude that for this  $u$ ,  $\mu_C(u) = f_{\vee}(p_1, \dots, p_j, \dots, p_K) \geq p_j > 0$ , so  $\mu_C(u)$  is not identically 0.

We now prove that if  $|u - g(\vec{x})| > \varepsilon$ , then  $\mu_C(u) = 0$  by showing that in this case  $p_1 = p_2 = \dots = p_j = \dots = p_K = 0$ , and so  $\mu_C(u) = f_{\vee}(p_1, p_2, \dots, p_K) = f_{\vee}(0, 0, \dots, 0) = 0$ .

Let us take an arbitrary  $j$  from the interval 1 to  $K$  and prove that  $p_j = 0$ . Indeed, since

$$p_j = f_{\&}(\mu_{j,1}(x_1), \mu_{j,2}(x_2), \dots, \mu_{j,n}(x_n), \mu_j(u)) > 0$$

and  $f_{\&}(0, p) = 0$ , the only possibility for  $p_j$  to be positive is for all the terms  $\mu_{j,1}(x_1), \mu_{j,2}(x_2), \dots, \mu_{j,n}(x_n)$ , and  $\mu_j(u)$  to be positive. Because of our choice of  $\mu_{j,i}$ , the first term is positive only for  $|x_1 - x_1^j| \leq \delta$ . In a similar manner, the first  $n$  terms are positive only if  $|x_i - x_i^j| \leq \delta$  for all  $i$ . In this case, by the choice of  $\delta$ ,  $|g(\vec{x}) - g(\vec{x}^j)| = |g(\vec{x}) - u^j| \leq \varepsilon/2$ , but we assumed that  $|u - g(\vec{x})| > \varepsilon$ ; therefore,  $|u - u^j| \geq |u - g(\vec{x})| - |g(\vec{x}) - u^j| > \varepsilon/2$ , and  $\mu_j(u) = 0$ .

For every  $\vec{x}$ , therefore, either one of terms  $\mu_{j,i}(x_i)$  is equal to 0, or they are all positive, in which case  $\mu_j(u) = 0$ . In both cases,  $p_j = 0$ , and so  $\mu_C(u) = f_{\vee}(p_1, \dots, p_K) = f_{\vee}(0, 0, \dots, 0) = 0$  for all  $u$  outside an interval  $[g(\vec{x}) - \varepsilon, g(\vec{x}) + \varepsilon]$ , and thence  $f(\vec{x}) = F(\mu_C)$  belongs to this same interval. Q.E.D.

**Proof of Theorem 3.** Since we assumed that  $\mathcal{M}$  is a universal fuzzy control tool, for every choice of values  $u_1^{(1)} \in U_1, \dots, u_1^{(N_1)} \in U_1, \dots, u_m^{(1)} \in U_m, \dots, u_m^{(N_m)} \in U_m$ , we can approximate an arbitrary continuous function

$$F(z_1, \dots, z_n, f_1(u_1^{(1)}), \dots, f_1(u_1^{(N_1)}), \dots, f_m(u_m^{(1)}), \dots, f_m(u_m^{(N_m)}))$$

by applying  $\mathcal{M}$  to an appropriate rule base. Therefore, to complete our proof, it is sufficient to prove that an arbitrary continuous functional  $J(s) = J(z_1, \dots, z_n, f_1, \dots, f_m)$  can be approximated by functions of the type

$$F(z_1, \dots, z_n, f_1(u_1^{(1)}), \dots, f_1(u_1^{(N_1)}), \dots, f_m(u_m^{(1)}), \dots, f_m(u_m^{(N_m)}))$$

for appropriately chosen values  $u_1^{(1)} \in U_1, \dots, u_1^{(N_1)} \in U_1, \dots, u_m^{(1)} \in U_m, \dots, u_m^{(N_m)} \in U_m$ . It turns out that this is indeed true if the values  $u_j^{(i)}$  form a  $\delta$ -net for sufficiently small  $\delta$ . This auxiliary result was proved in [11] (see the proof of Theorem 1 there; it is worth mentioning that this Theorem states that *neural networks* are a universal control tool for distributed systems).

To illustrate the approximation techniques, we present the proof for the simplest possible distributed system, viz., the case where its state is described by specifying only one function. Throughout this proof, we will use known mathematical results about compact sets (see, e.g., [18]).

Let  $F$  be a compact subset of  $C(U)$ , where  $U$  is a metric compact space, and let  $J : F \rightarrow R$  be continuous. Then for each  $\varepsilon > 0$  there exists a continuous function  $\pi$  from  $F$  into a finite-dimensional Euclidean space  $R^q$ , and a continuous function  $J_\varepsilon$  that is defined a compact subset  $\pi(F)$  of  $R^q$  such that  $|J(f) - J_\varepsilon(\pi(f))| \leq \varepsilon$  for every  $f \in F$ . Indeed:

- (i) Since  $F$  is compact,  $J$  is uniformly continuous on  $F$ , so, there exists  $\delta(\varepsilon)$  such that whenever  $\|f - g\| \leq \delta(\varepsilon)$ ,  $|J(f) - J(g)| \leq \varepsilon$ . Let  $G$  be the finite set of points in  $F$  such that for every  $f \in F$ , there exists a  $g \in G$  with  $\|f - g\| \leq \delta(\varepsilon)/3$  (the existence of such  $G$  also follows from the fact that  $F$  is a compact).
- (ii) Since  $F$  is a compact subset of  $C(U)$ ,  $F$  forms a family of equicontinuous functions, so that there exists a  $\beta(\varepsilon) > 0$  such that whenever  $\|u - v\| \leq \beta(\varepsilon)$ , we have  $|f(u) - f(v)| \leq \delta(\varepsilon)/4$  for every  $f \in F$ . Choose a finite set  $V \subseteq U$  such that for every  $u \in U$ , there exists a  $v \in V$  such that  $\|u - v\| \leq \beta(\varepsilon)$ ; we denote elements of  $V$  by  $v_1, \dots, v_q$ .
- (iii) Define  $\pi : F \rightarrow \pi(F) \subseteq R^q$  (where  $q$  is the cardinality of  $V$ ), by  $\pi(f) = (f(v_1), \dots, f(v_q))$ . Obviously,

$$\|\pi(f) - \pi(\tilde{f})\| = \max_{1 \leq i \leq q} |f(v_i) - \tilde{f}(v_i)| \leq \|f - \tilde{f}\|,$$

so  $\pi$  is continuous and hence  $\pi(F)$  is compact.

(iv) Define  $J_\varepsilon : \pi(F) \rightarrow R$  by

$$J_\varepsilon(\pi(f)) = \sum_G \alpha_g(f) J(g) / \sum_G \alpha_g(f),$$

where, for every  $g \in G$ ,  $\alpha_g(f) = \max\{0, \delta(\varepsilon)/2 - \|\pi(f) - \pi(g)\|\}$  is a continuous function of  $\pi(f)$ . By (i), for every  $f \in F$ , there exists a  $g \in G$  such that  $\|\pi(f) - \pi(g)\| \leq \|f - g\| \leq \delta(\varepsilon)/3 < \delta(\varepsilon)/2$ , so  $\sum_G \alpha_g(f) > 0$ . Thus,  $J_\varepsilon$  is well defined and continuous on  $\pi(F)$ .

(v) Observe that  $|J(f) - J_\varepsilon(\pi(f))| \leq \max |J(f) - J(g)|$ , where the maximum is taken over all  $g \in G$  such that  $\|\pi(f) - \pi(g)\| \leq \delta(\varepsilon)/2$ . Now, for every  $u \in U$ , there exists a  $v \in V$  such that  $\|u - v\| \leq \beta(\varepsilon)$ , so

$$\begin{aligned} |f(u) - g(u)| &\leq |f(u) - f(v)| + |f(v) - g(v)| + |g(v) - g(u)| \leq \\ &\leq \delta(\varepsilon)/4 + |f(v) - g(v)| + \delta(\varepsilon)/4. \end{aligned}$$

Hence,

$$\|f - g\| = \max_U |f(u) - g(u)| \leq \delta(\varepsilon)/2 + \|\pi(f) - \pi(g)\|.$$

Thus, when  $\|\pi(f) - \pi(g)\| \leq \delta(\varepsilon)/2$ , we have  $\|f - g\| \leq \delta(\varepsilon)$  and hence  $|J(f) - J(g)| \leq \varepsilon$ , implying that for every  $f \in F$ ,  $|J(f) - J_\varepsilon(\pi(f))| \leq \varepsilon$ . Q.E.D.

*Remark.* The above reduction from an infinite-dimensional case to a finite-dimensional case is primarily for practical implementation purposes. This procedure is parallel in spirit to, e.g., computational techniques in  $H^\infty$ -control (for definitions and details see, e.g., [4]), in which the  $H^\infty$ -optimization problem for distributed systems (which are infinite-dimensional) is reduced to the finite-dimensional problem of finding weighting filters.

**The proofs of Theorems 4 and 5** follow similarly from the results in [11].

## 9. CONCLUSIONS

The results of [2,7,20,21], as well as our own results, show that fuzzy control is a universal control tool in the following sense: for every (potential) control strategy and for every  $\varepsilon$  there exists a fuzzy controller whose control is  $\varepsilon$ -close to the desired one. In other words, even though we do not know what an optimal control is, we can in principle, by an appropriate choice of the rules, get a fuzzy control that is as close to the optimal control as possible. Thus, even if we restrict ourselves entirely to fuzzy control methodology, we still do not lose any possible control strategies.

These results were initially proved for plants whose states can be described by finitely many parameters. We proved in this paper two generalizations of this result, which show that fuzzy control can be used in the most general control situations, specifically, we proved the following two results:

- Fuzzy control is a universal control tool for distributed systems (i.e., plants that require infinitely many parameters to describe their current state).
- In principle, one can design a universal fuzzy controller: we input the description of a plant and the desired objective and generate an  $\varepsilon$ -optimal control.

In mathematical terms, we proved that fuzzy control systems can approximate arbitrary continuous functions and functionals.

These results are proved for an arbitrary fuzzy control methodology; the choice of a fuzzy methodology should be based, therefore, upon additional considerations such as smoothness, stability, and sensitivity of the resulting control (see, e.g., [8,9,16]).

**Acknowledgments.** This work was supported by a NSF Grant No. CDA-9015006, NSF Research Opportunity Award (for O.S.), NASA Research Grants No. NAG 9-482 and NAG 9-757, and a grant from the Institute for Materials and Manufacturing Management. The work of the first author was partially carried out at the Laboratory for International Fuzzy Engineering Research, Chair of Fuzzy Theory, Japan. We are thankful to the anonymous referees for their suggestions. We are especially thankful to Joseph M. Barone, guest editor of this special issue, who tirelessly went over our text phrase-by-phrase which helped us improve and clarify our paper.

## REFERENCES

- [1] H. R. Berenji, Fuzzy logic controllers, In: R. R. Yager, L. A. Zadeh, Eds., *An Introduction to Fuzzy Logic Applications in Intelligent Systems* (Kluwer Academic Publ., Boston, MA, 1991) 69–96.
- [2] J. J. Buckley, Sugeno type controllers are universal controllers, *Fuzzy Sets and Systems* 53 (1993) 299–303.
- [3] S. S. L. Chang and L. A. Zadeh. *On fuzzy mapping and control*, *IEEE Transactions on Systems, Man and Cybernetics SMC-2* (1972) 30–34.
- [4] C. Foias, Commutant lifting techniques for computing optimal  $H^\infty$  controllers, In:  *$H^\infty$ -control theory*, Lecture Notes in Mathematics, Vol. 1496 (Springer-Verlag, N.Y., 1991) 1–36.
- [5] I. R. Goodman and H. T. Nguyen. *Uncertainty models for knowledge-based systems* (North Holland, Amsterdam, 1985).
- [6] K. Hornik, M. Stinchcombe, and H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* 2 (1989) 359–366.
- [7] B. Kosko, Fuzzy systems as universal approximators, *Proceedings of the 1st IEEE International Conference on Fuzzy Systems*, San Diego, CA, 1992, 1153–1162.
- [8] V. Kreinovich, C. Quintana, and R. Lea, What procedure to choose while designing a fuzzy control? Towards mathematical foundations of fuzzy control, *Working Notes of the 1st International Workshop on Industrial Applications of Fuzzy Control and Intelligent Systems*, College Station, TX, 1991, 123–130.

- [9] V. Kreinovich, C. Quintana, R. Lea, O. Fuentes, A. Lokshin, S. Kumar, I. Boricheva, and L. Reznik, What non-linearity to choose? Mathematical foundations of fuzzy control, *Proceedings of the 1992 International Fuzzy Systems and Intelligent Control Conference*, Louisville, KY, 1992, 349–412.
- [10] V. Kreinovich, C. Quintana, L. Reznik, Gaussian membership functions are most adequate in representing uncertainty in measurements, *Proceedings of NAFIPS'92: North American Fuzzy Information Processing Society Conference, Puerto Vallarta, Mexico, December 15–17, 1992*, NASA Johnson Space Center, Houston, TX, 1992, Vol. II, 618–624.
- [11] V. Kreinovich and O. Sirisaengtaksin, 3-layer neural networks are universal approximators for functionals and for control strategies, *Neural, Parallel, and Scientific Computations* 1 (1993) 325–346.
- [12] C. C. Lee, Fuzzy logic in control systems: Fuzzy logic controllers, parts I and II, *IEEE Transactions on Systems, Man and Cybernetics* 20 (1990) 404–435.
- [13] E. H. Mamdani, Application of fuzzy algorithms for control of simple dynamic plant, *Proceedings of the IEE* 121, No. 12 (1974) 1585–1588.
- [14] R. R. Mohler, *Nonlinear systems. Vol. 1. Dynamics and control* (Prentice Hall, Englewood Cliff, NJ, 1991).
- [15] H. T. Nguyen and V. Kreinovich, On approximation of controls by fuzzy systems, *Proceedings of the Fifth International Fuzzy Systems Association World Congress*, Seoul, Korea, July 1993, 1414–1417.
- [16] H. T. Nguyen, V. Kreinovich, and D. Tolbert, On robustness of fuzzy logics, *Proceedings of the 2nd IEEE-FUZZ International Conference*, San Francisco, CA, March 1993, Vol. 1, 543–547.
- [17] G. C. Oden, Integration of fuzzy logical information, *Journal of Experimental Psychology: Human Perception Perform.*, 3, No. 4 (1977) 565–575.
- [18] D. A. Sprecher, *Elements of real analysis* (Dover Publ., N.Y., 1987).
- [19] M. Sugeno (Ed.), *Industrial applications of fuzzy control* (North Holland, Amsterdam, 1985).
- [20] L.-X. Wang, Fuzzy systems are universal approximators, *Proceedings of the 1992 IEEE International Conference on Fuzzy Systems*, San Diego, CA, 1992, 1163–1169.
- [21] L.-X. Wang and J. Mendel, *Generating fuzzy rules from numerical data, with applications*, University of Southern California, Signal and Image Processing Institute, Technical Report USC-SIPI # 169 (1991).
- [22] L. Zadeh, Fuzzy sets, *Information and control* 8 (1965) 338–353.
- [23] H. J. Zimmermann, Results of empirical studies in fuzzy set theory, In: G. J. Klir, Ed., *Applied General System Research* (Plenum, N.Y., 1978) 303–312.