Adding Fuzzy Integral to Fuzzy Control

H.T. Nguyen
Department of
Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003, USA
hunguyen@nmsu.edu

V. Kreinovich
Department of
Computer Science
University of Texas
El Paso, TX 79968, USA
vladik@cs.utep.edu

R. Aló
Department of Computer &
Mathematical Sciences
Univ. of Houston-Downtown
Houston, TX 77002, USA
ralo@uh.edu

Abstract

Sugeno integral was invented a few decades ago as a natural fuzzy-number analogue of the classical integral. Sugeno integral has many interesting applications. It is reasonable to expect that it can be used in all application areas where classical integrals are used, and in many such areas it is indeed useful. Surprisingly, however, it has never been used in *fuzzy control*, although in traditional control, classical integral is one of the main tools.

In this paper, we show that the appropriately modified Sugeno integral is indeed useful for fuzzy control: namely, it provides numerical characterization of stability and smoothness of fuzzy control strategies.

1 Introduction

1.1 The problem

Two decades ago, Sugeno discovered a natural fuzzy analogue of the classical integral. Sugeno integral has many interesting applications.

It is reasonable to expect seems that Sugeno integral can be used in all application areas where classical integrals are used, and in many such areas it is indeed useful. Surprisingly, however, it has never been used in *fuzzy control*, although in traditional control, classical integral is one of the main tools.

It is even more surprising because fuzzy control is one of the main applications of fuzzy systems theory, probably its most deeply-researched and widely applied area.

1.2 What we are planning to do

In this paper, we explain why the *original* Sugeno integral has not been used in fuzzy control, and how we can *modify* this notion so that it will lead to useful applications: namely, it will provide numerical characterization of stability and smoothness of fuzzy control strategies.

1.3 The structure of the paper

In Section II, we start with a brief reminder of the notion of Sugeno integral. Since our main goal is to *modify* this definition, we will give detailed motivations for the original Sugeno's definition, motivations that will help us understand:

- why this definition is not used in fuzzy control, and
- how we can modify it so that it becomes applicable.

In Section III, we describe the numerical characteristics of the desired properties of the resulting control, properties such as stability and smoothness. Finally, in Section IV, we will show how an appropriate modification of Sugeno integral can describe these characteristics for fuzzy control.

Preliminary results of this research appeared in [5].

2 Sugeno integral: a brief reminder

2.1 Classical Integral as a Natural Method of Handling Real Numbers That Describe the Physical World

Why are we interested in the classical integral? Sugeno integral is a natural fuzzy analogue of the classical integral. Therefore, to describe the ideas behind Sugeno integral, we must briefly recall where the classical integral came from.

Addition as a natural operation with real numbers describing the physical world. Classical mathematics of real numbers was originally invented to describe the values of physical quantities. For physical quantities, the most fundamental operations are those that have direct physical meaning:

- *addition* $x+y$, which corresponds, e.g. to the case when we combine two bodies together; then, the mass and the charge of the combination is equal to the sum of the masses (charges) of the combined component bodies;
- *multiplication*, which corresponds, e.g., to *re-scaling* of physical quantities, etc.

Sums of many numbers, and integral as a description of such sums. In many physical problems, we combine *finitely* many quantities and therefore, it is sufficient to consider finite sums and products.

However, in some other practically important physical situations, the number of combined quantities is so large that it is reasonable to assume that we are combining *infinitely many* quantities. For example, a typical macro-size body consists of $\approx 10^{23}$ molecules. This amount is so huge that even with the modern computers, it is absolutely not possible to handle these molecules individually. No matter into how many pieces our computer model divides the body, each piece still contains many molecules and can be, therefore, divided even further. In other words, a multi-molecule macro-size body behaves, in our modeling, as if there is no limit to division, i.e., as if we actually have *infinitely many* parts. This *continuous* approximation is indeed very useful in physics.

In this approximation (i.e., in this limit), the *sum* turns into its limit, which is exactly the classical *integral*.

2.2 Handling expert information: a modified integral is needed

Why fuzzy? Traditional (pre-fuzzy) mathematics is very good in processing measurement results. In some applications areas, this is quite sufficient for making reasonable decisions, but in many application areas, measurement results themselves are not sufficient for making reasonable decisions: In addition to the measurement results, we must take into consideration human expertise, expertise which is usually formulated not in precise mathematical terms, but by words from natural language.

To formalize this expert knowledge (and thus, to make it available for computers and therefore, for automation), L. Zadeh invented *fuzzy logic* and *fuzzy set theory* [12].

Enter real numbers. In the original version of fuzzy set theory, *real numbers* were used to describe the expert's degrees of certainty (also known as *truth values*). Later on, more general objects have been proposed to describe truth values, but so far, real numbers from the interval $[0,1]$ remain the main truth value domain in applications of fuzzy logic and fuzzy set theory.

In fuzzy logic, real numbers are the same as in crisp data processing, but natural operations are different. Since we use the same real numbers as in traditional mathematics, at first glance, it may seem that all notions of classical mathematics, including the notion of an integral, are naturally applicable to fuzzy values as well. In reality, the situations is not so straightforward:

- Indeed, for real numbers that describe the values of physical quantities, the most fundamental operation is addition, and therefore, an integral, which is, in essence, nothing else but a limit of the sums, is a very natural tool.
- On the other hand, for real numbers that describe the expert information, addition is an artificial operation, which not only is not easy to interpret, but that is even not always applicable: e.g., since we are only considering real numbers from the interval $[0,1]$, sums like $1+1$ whose values are greater than 1 have no meaning at all. Since the *sum* is not natural, the *integral* is not natural either.

We do need a fuzzy analogue of the integral. This un-naturalness does not mean that we can simply dismiss the notion of an integral from fuzzy logic: Similarly to processing measurement results, we often face the necessity of processing many expert estimates, and it would be nice to have an integral-like limit tool that can handle these situations. Thus, it is desirable to develop *fuzzy analogues* of the integral, analogues that are based, instead of addition and multiplication, on operations that are more natural in fuzzy logic.

Operations with real numbers that are natural in fuzzy logic. Since fuzzy logic describes human reasoning, the most natural operations in fuzzy logic are "or" and "and" operations $\vee$ and $\&$ on truth values (these operations are usually called t-conorms and t-norms; see, e.g., [2, 8]).

There is an analogy between logical and arithmetic operations that goes back to Boole, the founding father of modern formalized logic:

- $\vee$ is a natural analogue of addition, and
- $\&$ is a natural analogue of multiplication.

In fuzzy logic, this analogy goes even further than in classical logic:

- multiplication is actually one of the most widely used version of an "and" operation (it is one of the two "and" operations proposed by L. Zadeh in his pioneer paper), and
- one of the widely used "or" operations is $a \vee b = \min(a+b, 1)$, i.e., addition corrected in such a way that the result of this operation always stays within an integral $[0,1]$.

It is therefore, desirable to describe an analogue of the classical interval in which addition is replaced by an "or" operation and multiplication by an "and" operation.

This idea was first implemented by M. Sugeno in his dissertation [10]; the main results of this dissertation were later published in a paper form [11]; for a latest survey on Sugeno integral and related topics, see, e.g., [1].

2.3 Sugeno integral: deriving the original definition

The domain of the future definition. We have already mentioned that in physical applications, the classical integral is used to describe the case when we have a large number of variables $x_1, \ldots, x_n$, and when, instead of describing them separately, we assume that we have a *continuous* family of variables $x(t)$ depending on a continuous parameter t. Then, the sum $\sum x_i$ tends to a limit $\int x(t)\,dt$.

Similarly, in fuzzy case, we consider the situation, when we have expert degree of certainty μ_i assigned to a large number of situations. So, instead of analyzing these situations one-by-one, we assume that we actually have a degree of certainty $\mu(t)$ for an *arbitrary* value of the continuously changing parameter. This assumption goes back to the pioneer paper of Zadeh: it represents a *membership function* depending on the parameter t. So, the fuzzy analogue of the classical integral should be applicable to membership functions.

First try. Historically, an integral $\int x(t)\,dt$ was first defined as a limit of the integral sums $\sum x(t_i) \cdot \Delta t$. A natural idea is, therefore, to replace the sum by a "or" operation $\vee$, i.e., to consider $\vee_t(\mu(t) \cdot \Delta t)$, where $\vee_t$ means that $\vee$ is applied to infinitely many values $\mu(t)$ of the membership function $\mu(t)$. In particular, if we take max as an "or" operation, we have $(\max_t \mu(t)) \cdot \Delta t$.

Drawback of the resulting definition. The main drawback of the resulting definition is that most membership functions are *normalized*, i.e., for them $\mu(t) = 1$ for some t. For such membership functions, the above-defined "integral" is simply equal to Δt.

Why this drawback? To find out why this definition did not work, let us go back to the original definition of the integral that we were trying to fuzzify. This definition was originally proposed and used for *continuous* functions.

- For functions that describe the dependency between *physical quantities*, *continuity* is a very natural requirement: For example, if we know that voltage is determined by current, this means that small changes in voltage should lead to small changes in current, i.e., that the dependence should be *continuous*.
- On the other hand, for *membership functions*, continuity is not so natural. A reasonable example of a membership function is an example of a *crisp* property, in which for every t, we are either sure that this property is true for t (i.e., $\mu(t) = 1$), or we are sure that this property is false for t, i.e., $\mu(t) = 0$. Such functions, that only take values 0 and 1, are *not continuous*.

For discontinuous functions, the above definition of an integral often does not work, so no wonder that its fuzzy version is not working either.

To make it work, we need to fuzzify a *different* definition of the classical integral, a definition that would be applicable to discontinuous functions as well.

Sugeno's definition of an integral. The extension of the notion of the integral to complicated discontinuous functions was proposed by Lebesgue, the founder of the modern integration theory. For his generalization, Lebesgue used the fact that the function $x(t)$ itself can be represented as an integral $\int_0^{x(t)} 1\,d\alpha$. If we substitute this formula into the desired integral $\int x(t)\,dt$, and swap the variables t and α, we conclude that $\int x(t)\,dt = \int \mu_0(\{t \,|\, x(t) \geq \alpha\})\,d\alpha$, where $\mu_0(A)$ denotes the Lebesgue measure of a set A (we can take any other measure instead of Lebesgue's measure). If we replace the product by "and" and the sum by "or", we get Sugeno's formula $\vee_\alpha(\alpha \& \mu_0(\{t \,|\, \mu(t) \geq \alpha\}))$.

Usually, in this formula, the simplest possible "and" and "or" operations are taken: $\vee = \max$ and $\& = \min$. In this case, Sugeno's formula turns into
$\max_\alpha \min(\alpha, \mu_0(\{t \,|\, \mu(t) \geq \alpha\}))$.

2.4 Successes and drawbacks of the original Sugeno's definition

Successes. Sugeno's definition is very suitable for describing expert knowledge. Let us give an example. One of the most natural way to assign the values to a membership function $\mu(t)$ that describes a certain property P is to use *polling*: for every value t, we ask experts whether they believe that t satisfies this property P, and take, as $\mu(t)$, the fraction of experts who answered "yes".

In some cases, the property P is true for all values t. In other words, for 100% of all the values t, 100% of all the experts believe that P is true (i.e., $\mu(t) = 1$). What does it means that a property is, say, at least 90% true? It is natural to define this notion by requiring that for at least 0.9 of all values t, at least 90% of all experts believe that t is true. In other words, for some $\alpha \geq 0.9$, we have $\mu_0(\{t \,|\, \mu(t) \geq \alpha\}) \geq 0.9$, i.e., we have $\max_\alpha \min(\alpha, \mu_0(\{t \,|\, \mu(t) \geq \alpha\})) \geq 0.9$. Thus, the degree d to which all elements satisfy the property P can be defined as the largest d which satisfies this inequality for some α, i.e., as one can easily check, as the Sugeno integral.

Sugeno integral is also efficiently used in image processing and in other important computer applications areas.

Main drawback. There is only one area where fuzzy logic is actively used but where applications of Sugeno integral are lacking: fuzzy control. At first glance, this is very strange, for three reasons:

- Fuzzy control is, currently, one of the main applications of fuzzy logic, and probably the most well-researched one.
- In traditional control, integration is one of the main tools, so one should expect that a fuzzy analogue of the classical integral should be widely in fuzzy control as well.
- Sugeno himself is not only the author of the notion of Sugeno's integral, but he is one of the world's leading researchers in theory and application of fuzzy control, and still he has not found a relationship between these two areas in which he is so proficient.

In view of this three reasons, the absence of applications cannot be attributed simply to lack of trying, we rather view it as a drawback of the original Sugeno's definition.

Why this drawback, and what we can do about it. The above motivations for Sugeno integral explain why the existing form of Sugeno integral is not directly used in fuzzy control:

- Traditional integral is based on the addition and multiplication operations that are natural for measurement results but un-natural for expert degrees of certainty. Therefore, this integral is very useful when we only have measurement results and no expert information (e.g., in traditional control).
- Sugeno integral is based on the operations $\vee$ and & that are natural for expert degrees of uncertainty, but un-natural for measurement results. It is therefore very useful in the situations when we only have expert information but few measurement results (e.g., in traditional expert systems).
- In fuzzy control, however, we need *both* the measurement results *and* the expert estimates. So, we cannot use the original formula for Sugeno integral, because the operations underlying this formula are un-natural for half of the data.

In view of this conclusion, what we need for fuzzy control is a *modification* of Sugeno integral that uses both arithmetic operations (addition and multiplication) that are natural for measurement "half" of the data and logical operations that are natural for the expert "half" of the data.

In the next two sections, we will see that this modification is indeed possible and helpful. Before we start doing that, let us describe which numerical characteristics of fuzzy control it is desirable to describe.

3 Numerical characteristics of the ideal control

3.1 What is ideal control?

Engineers rarely explain explicitly what exactly they mean by an *ideal* control. However, they often do not hesitate to say that one control is better than another one. What do they mean by that? Usually, they draw a graph that describes how an initial perturbation changes with time, and they say that a control is good if this perturbation quickly goes down to 0 and then stays there.

In other words, an ideal control consists of two stages:

- On the *first stage*, the main objective is to make the difference $x = X - X_0$ between the actual state X of the plant and its ideal state X_0 go to 0 as fast as possible.
- After we have already achieved the objective of the first stage, and the difference is close to 0, then the *second stage* starts. On this second stage, the main objective is to keep this difference close to 0 at all times. We do not want this difference to oscillate wildly, we want the dependency $x(t)$ to be as smooth as possible.

This description enables us to formulate the objectives of each stage in precise mathematical terms.

3.2 First stage of the ideal control: main objective

For readers' convenience, we will illustrate our ideas on a simple plant. So, let us consider the case when the state of the plant is described by a single variable x, and we control the first time derivative $\dot{x}$. For this case, we arrive at the following definition:

Definition 1. *Let a function $u(x)$ be given (this function will be called a control strategy). By a trajectory of the plant, we understand the solution of the differential equation $\dot{x} = u(x)$. Let's fix a positive number M (e.g., $M = 1000$). Assume also that a real number $\delta \neq 0$ is given. This number will be called an initial perturbation. A relaxation time $t(\delta)$ for the control $u(x)$ and the initial perturbation δ is defined as follows:*

- *we find a trajectory $x(t)$ of the plant with the initial condition $x(0) = \delta$, and*
- *we take as $t(\delta)$, the first moment of time starting from which $|x(t)| \leq |x(0)|/M$ (i.e., for which this inequality is true for all $t \geq t(\delta)$).*

For *linear control*, i.e., when $u(x) = -kx$ for some constant k, we have $x(t) = x(0)\exp(-kt)$ and therefore, the relaxation time t is easily determined by the equation $\exp(-kt) = 1/M$, i.e., $t = \ln(M/k)$. Thus defined relaxation time does not depend on δ. So, for control strategies that use linear control on the first stage, we can easily formulate the objective: to minimize relaxation time. The smaller the relaxation time, the closer our control to the ideal.

In the *general case*, we would also like to minimize relaxation time. However, in general, we encounter the following problem: For *non-linear control* (and fuzzy control *is* non-linear) the relaxation time $t(\delta)$ depends on δ. If we pick a δ and minimize $t(\delta)$, then we get good relaxation for this particular δ, but possibly at the expense of not-so-ideal behavior for different values of the initial perturbation δ.

What to do? The problem that we encountered was due to the fact that we considered a simplified control situation, when we start to control a system only when it is already out of control. This may be too late. Usually, no matter how smart the control is, if a perturbation is large enough, the plant will never stabilize. For example, if the currents that go through an electronic system exceed a certain level, they will simply burn the electronic components. To avoid that, we control the plant from the very beginning, thus preventing the values of x from becoming too large. From this viewpoint, what matters is how fast we go down for *small* perturbations, when $\delta \approx 0$.

What does "small" mean in this definition? If for some value δ that we initially thought to be small, we do not get a good relaxation time, then we will try to keep the perturbations below that level. On the other hand, the smaller the interval that we want to keep the system in, the more complicated and costly this control becomes. So, we would not decrease the admissible level of perturbations unless we get a really big increase in relaxation time. In other words, we decrease this level (say, from δ_0 to $\delta_1 < \delta_0$) only if going from $t(\delta_0)$ to $t(\delta_1)$ means decreasing the relaxation time. As soon as $t(\delta_1) \approx t(\delta_0)$ for all $\delta_1 < \delta_0$, we can use δ_0 as a reasonable upper level for perturbations.

In mathematical terms, this condition means that $t(\delta_0)$ is close to the limit of $t(\delta)$ when $\delta \to 0$. So, the smaller this limit, the faster the system relaxes. Therefore, this limit can be viewed as a reasonable objective for the first stage of the control.

Definition 2. *By a relaxation time T for a control $u(x)$, we mean the limit of $t(\delta)$ for $\delta \to 0$.*

So, *the main objective of the first stage of control is to maximize relaxation time.*

3.3 Second stage of the ideal control: main objective

After we have made the difference x go to 0, the second stage starts, on which $x(t)$ has to be kept as smooth as possible. What does *smooth* mean in mathematical terms? Usually, we say that a trajectory $x(t)$ is smooth at a given moment of time t_0 if the value of the time derivative $\dot{x}(t_0)$ is close to 0. We want to say that a trajectory is smooth if $\dot{x}(t)$ is close to 0 for all t.

In other words, if we are looking for a control that is the smoothest possible, then we must find the control strategy for which $\dot{x}(t) \approx 0$ for all t. There are infinitely many moments of time, so even if we restrict ourselves to control strategies that depend on finitely many parameters, we will have infinitely many equations to determine these parameters. In other words, we will have an *over-determined* system. Such situations are well-known in data processing, where we often have to find parameters $p_1, \ldots, p_n$ from an over-determined system $f_i(p_1, \ldots, p_n) \approx q_i, 1 \le i \le N$. A well-known way to handle such situations is to use the *least squares method*, i.e., to find the values of p_j for which the "average" deviation between f_i and q_i is the smallest possible. To be more precise, we minimize the sum of the squares of the deviations, i.e., we are solving the following minimization problem:

$$\sum_{i=1}^{N} \left(f_i(p_1, \ldots, p_n) - q_i)^2\right) \to \min_{p_1,\ldots,p_n}.$$

In our case, $f_i = \dot{x}(t)$ for different moments of time t, and $q_i = 0$. So, least squares method leads to the criterion $\sum(\dot{x}(t))^2 \to \min$. Since there are infinitely many moments of time, the sum turns into an integral, and the criterion for choosing a control into $J(x(t)) \to \min$, where $J(x(t)) = \int (\dot{x}(t))^2\, dt$. This value J thus represents a degree to which a given trajectory $x(t)$ is non-smooth. So, we arrive at the following definition:

Definition 3. *Assume that a control strategy $x(t)$ is given, and an initial perturbation δ is given. By a non-smoothness $I(\delta)$ of a resulting trajectory $x(t)$, we understand the value $J(x) = \int_0^\infty (\dot{x}(t))^2\, dt$.*

The least squares method is not only heuristic, it has several reasonable justifications. So, instead of simply borrowing the known methodology from data processing (as we did), we can formulate reasonable conditions for a functional J (that describes non-smoothness), and thus deduce the above-described form of J without using analogies at all. This is done in [4].

What control to choose on the second stage? Similarly to relaxation time, we get different criteria for choosing a control if we use values of non-smoothness that correspond to different δ. And similarly to relaxation time, a reasonable solution to this problem is to choose a control strategy for which in the limit $\delta \to 0$, the non-smoothness takes the smallest possible value.

Mathematically, this solution is a little bit more difficult to implement than the solution for the first stage: Indeed, the relaxation time $t(\delta)$ has a well-defined non-zero limit when $\delta \to 0$, while non-smoothness simply tends to 0. Actually, for linear control, $I(\delta)$ tends to 0 as δ^2. To overcome this difficulty and still get a meaningful limit of non-smoothness, we will divide $J(x)$ (and, correspondingly, $I(\delta)$) by δ^2 and only then, tend this ratio $\tilde{J}(x(t)) = \tilde{I}(\delta)$ to a limit. This division does not change the relationship between the functional and smoothness: indeed, if for some δ, a trajectory $x_1(t)$ is smoother than a trajectory $x_2(t)$ in the sense that $J(x_1(t)) < J(x_2(t))$, then, after dividing both sides by δ^2, we will get $\tilde{J}(x_1(t)) < \tilde{J}(x_2(t))$. So, a trajectory $x(t)$ for which $\tilde{J}(x)$ is smaller, is thus smoother.

As a result, we arrive at the following definition.

Definition 4. *By a non-smoothness I of a control $u(x)$, we mean the limit of $I(\delta)/\delta^2$ for $\delta \to 0$.*

Thus, *the main objective of the second stage of control is to minimize non-smoothness.*

4 Modified Sugeno integral helps

4.1 Fuzzy control: in brief

In general, fuzzy control starts with the rules of the 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 x_i are parameters that characterize the plant, u is the control, and A_i^j, B^j are the terms of natural language that are used in describing j–th rule (e.g., "small", "medium", etc).

The value u is a proper value of the control if and only if one of these rules is applicable. Therefore, if we use the standard mathematical notations & for "and", $\vee$ for "or", and $\equiv$ for "if and only if", then the property "u is a proper control for a given x" (which we will denote by $C(u, x)$) can be described by the following informal "formula":

$$C(u,x) \equiv (A_1^1(x_1) \,\&\, A_2^1(x_2) \,\&\, \ldots \,\&\, A_n^1(x_n) \,\&\, B^1(u)) \vee$$
$$(A_1^2(x_1) \,\&\, A_2^2(x_2) \,\&\, \ldots \,\&\, A_n^2(x_n) \,\&\, B^2(u)) \vee$$
$$\ldots$$
$$(A_1^K(x_1) \,\&\, A_2^K(x_2) \,\&\, \ldots \,\&\, A_n^K(x_n) \,\&\, B^K(u))$$

If we use membership functions to describe these natural-language terms, we describe $A_i^j(x)$ as $\mu_{j,i}(x)$, the degree to which given value x satisfies the property A_i^j. Similarly, $B^j(u)$ is represented as $\mu_j(u)$. Then, after choosing an appropriate & and $\vee$ operations, we get the membership function for control: $\mu_C(u, x) = f_\vee(p_1, \ldots, p_K)$, where

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

To get a unique control value, in this paper, we will use a *centroid defuzzification*

$$\bar{u}(x) = \frac{\int u \cdot \mu_C(u,x)\,du}{\int \mu_C(u,x)\,du}.$$

For detailed description and alternatives, see, e.g., [3, 7].

4.2 Case study: a simple plant

Plant. In this paper, we will consider a simple system, in which the state of the plant is described by a single variable x, and we control its time derivative $\dot{x}$.

Properties and membership functions. Both for x and for $u = \dot{x}$, we will consider the membership functions that are most frequently used in fuzzy control applications:

- A property $M_0(x)$ ("x is negligible") is characterized by an even function $\mu_0(x)$ ($\mu_0(x) = \mu_0(-x)$) that is different from 0 only on the interval $(-\Delta, \Delta)$ and for which $\mu_0(1) = 1$;
- We also have properties $M_i(x)$ described by functions $\mu_i(x) = \mu_0(x - i \cdot \Delta)$. For example, $M_1(x)$ describes "small positive", the property $M_{-1}(x)$ describes "small negative", etc.

Membership functions that describe the properties $N_i(u)$ of the control u are assume to have the same shape. These functions may differ by a scaling, but, without losing generality, we can always assume that the units for both x and u are chosen in such a way that the membership functions for x and u are simply identical.

In the above definitions of stability and smoothness, we used *derivatives*. Usually, in applications, *continuous* membership functions are considered which are not necessarily differentiable. It is known, however, that an arbitrary continuous function can be approximate, with an arbitrary given accuracy, by a differentiable function. Therefore, without losing generality, we will assume that the function $\mu_0(x)$, when limited to the interval $[0, \Delta]$, is everywhere differentiable; we will also assume that, like for a triangular function, the derivative $\mu_0'(\Delta)$ at $x = \Delta$ is different from 0.

Rules. According to our definitions of stability and smoothness, we are only interested in the values of the x that are close to 0. For such values, only three of the above membership functions may be different from 0: M_0, M_1, and M_{-1}. Moreover, for every x, only two membership functions are different from 0:

- when $x < 0$, we only need M_0 and M_{-1};
- when $x > 0$, we only need M_0 and M_1.

For such x, the only reasonable rules are: "if $M_0(x)$ then $N_0(u)$", "if $M_{-1}(x)$ then $N_1(u)$", and "if $M_1(x)$ then $N_{-1}(u)$".

Choice of "and" and "or" operations. In principle, different "and" and "or" operations are used in fuzzy control. Some of these operations are not everywhere differentiable, which, for us, is a drawback, because we want the resulting formulas to be differentiable. To resolve this problem, we can use our recent result [6] that an arbitrary continuous t-norm can be, with any given accuracy, approximated by a *strictly Archimedean* t-norm, i.e., by a t-norm of the type $a\&b = \psi^{-1}(\psi(a)\cdot\psi(b))$, for some continuous function $\psi(a)$. The continuous function $\psi(a)$ can, in turn, be approximated by a *smooth* one (and we can also select this smooth approximation in such a way that $\psi'(0) \neq 0$). Thus, an arbitrary continuous t-norm, can be approximated, with an arbitrary accuracy, by a strictly Archimedean t-norm with a smooth function $\psi(a)$. Therefore, without losing generality, we can assume that out t-norm has this form.

Similarly, without losing generality, we can assume that out t-conorm ("or" operation) has the form $a\vee b = \varphi^{-1}(\varphi(a)+\varphi(b))$ for some differentiable function $\varphi(a)$ for which $\varphi'(0) \neq 0$.

4.3 Results

Preliminary results. The above definitions of characteristics of stability and smoothness were rather complicated, so it was not clear how to compute them. It is, therefore, desirable to re-formulate these definition in more directly computable terms. This reformulation is given in [4, 9]. Namely, if the resulting control strategy $\bar{u}(x)$ is a differentiable function x, then the *relaxation time* is equal to $\ln(M)/(-\bar{u}'(0))$, and the *non-smoothness* is equal to $I = -1/(2\bar{u}'(0))$. (The proof follows from the fact that for small x, $\bar{u}(x) \approx \bar{u}'(0)\cdot x$.) Thus, to compute both characteristics, we must estimate $|\bar{u}'(0)|$.

Derivation: main ideas. We are using the centroid defuzzification formula, according to which $\bar{u}(x) = N(x)/D(x)$, where $N(x) = \int u\cdot\mu_C(u,x)\,du$ and $D(x) = \int \mu_C(u,x)\,du$. For $x=0$, we get $\mu_C(u,0) = \mu_0(u)$. Since the function $\mu_0(u)$ is even, we have $N(0)=0$. Thus, we can conclude that the derivative $\bar{u}'(0)$ of the fraction $N(x)/D(x)$ is equal to $N'(0)/D(0)$.

Let us compute $N'(0)$ and $D(0)$. From $\mu_C(u,0) = \mu_0(u)$, we conclude that $D(0) = \int \mu_0(u)\,du$, and, since the function $\mu_0(u)$ is even, we conclude that $D(0) = 2\int_0^\Delta \mu_0(u)\,du$.

Let us now compute $N'(0)$, i.e., equivalently, the linear term $N'(0)\cdot x$ in the expression $N(x) = N'(0)\cdot x + o(x)$ for small negative x. When x is negative small ($|x| \le \Delta$), only two properties of x are satisfied with non-zero degree: "x is negligible" with the membership function $\mu_0(x) = 1+\mu_0'(0)\cdot x + o(x)$, and "$x$ is small negative", with the membership function $\mu_{-1}(x) = \mu_0(x+\Delta) = 0 + \mu_0'(\Delta)\cdot x + o(x)$. For these two membership functions, only two rules are applicable: "if $M_0(x)$ then $N_0(u)$" (we will denote it by R_1) and "if $M_{-1}(x)$ then $N_1(u)$" (we will denote this rule by R_2). The corresponding membership functions for u are different from 0 only when $u \in [-\Delta, 2\Delta]$, so the resulting function $\mu_C(u,x)$ can only be different from 0 for such values u. We can, therefore, represent the integral $N(x)$ as the sum of two integrals $N_1(x) = \int_{-\Delta}^{\Delta} u\cdot\mu_C(u,x)\,dx$ and $N_2(x) = \int_\Delta^{2\Delta} u\cdot\mu_C(u,x)\,dx$.

Let us start with the integral $N_2(x)$. When $u \in [\Delta, 2\Delta]$, only the rule R_2 is applicable. For this rule,

$$\mu_C(u,x) = p_2 = f_\&(\mu_{-1}(x), \mu_1(u)) =$$
$$f_\&(\mu_0'(\Delta)\cdot x + o(x), \mu_0(u-\Delta)).$$

When one of the arguments of the "and"-operation is close to 0, we get

$$f_\&(z,t) = \psi^{-1}(\psi(z)\cdot\psi(t)) =$$
$$\psi^{-1}(\psi'(0)\cdot z\cdot\psi(t)+o(z)) = (\psi^{-1})'(0)\cdot\psi'(0)\cdot z\cdot\psi(t)+o(z).$$

Using the formula for the derivative of the inverse function, we conclude that $f_\&(z,t) = z\cdot\psi(t)+o(z)$. Thus, in our case,

$$\mu_C(u,x) = p_2 = \mu_0'(\Delta)\cdot x\cdot\psi(\mu_0(u-\Delta)) + o(x).$$

Hence, the corresponding integral $N_2(x)$ can be represented as $x\cdot N_2'(0)+o(x)$, where

$$N_2'(0) = \mu_0'(\Delta)\cdot\int_\Delta^{2\Delta} u\cdot\psi(\mu_0(u-\Delta))\,du.$$

It is beneficial to use a new variable $v = u-\Delta$ which runs from 0 to Δ. In terms of this variable, $N_2'(0) = \mu_0'(\Delta)\cdot\int_0^\Delta (v+\Delta)\cdot\psi(\mu_0(v))\,dv$.

Let us now compute $N_1'(0)$. The integral $N_1(x)$ is obtained when u goes from $-\Delta$ to Δ. Alternatively, we can say that u goes from 0 to Δ, but for each such u, we can add the values which correspond to u and $-u$, i.e., we can represent $N_1(x)$ as

$$\int_0^\Delta [u\cdot\mu_C(u,x) + (-u)\cdot\mu_C(-u,x)]\,du =$$
$$\int_0^\Delta u\cdot[\mu_C(u,x) - \mu_C(-u,x)]\,du.$$

Let us find the explicit expressions for the integrand. For $-u \in [-\Delta, 0]$, only the rule R_1 is applicable; so, $\mu_C(-u,x) = p_1 = f_\&(\mu_0(-u), \mu_0(x))$. For small x, $\mu_0(x) = \mu_0(1) + x\cdot\mu_0'(0) + o(x)$. Here, $\mu_0(1) = 1$. Since μ_0 is an even function, we have $\mu_0(-u) = \mu_0(u)$ and $\mu_0'(1) = 0$; hence,

$$\mu_C(-u,x) = p_1 = f_\&(\mu_0(u), 1+o(x)) =$$
$$f_\&(\mu_0(u), 1) + o(x) = \mu_0(u) + o(x).$$

When $u \in [0,\Delta]$, both rules R_1 and R_2 are applicable, and so, $\mu_C(u,x) = f_\vee(p_1,p_2)$. We know that $p_2 = O(x)$, so we can conclude that $f_\vee(p_1,p_2) = \varphi^{-1}(\varphi(p_1)+\varphi(p_2)) = \varphi^{-1}(\varphi(p_1)+\varphi'(0)\cdot p_2 + o(p_2)) = \varphi^{-1}(\varphi(p_1)) + (\varphi^{-1})'(\varphi(p_1))\cdot p_2 + o(p_2)$. By definition of the inverse function, $\varphi^{-1}(\varphi(p_1)) = p_1$. Due to the formula of the derivative of the inverse function, we conclude that $(\varphi^{-1})'(\varphi(p_1)) = 1/\varphi'(p_1)$. Thus, we have $f_\vee(p_1,p_2) = p_1 + \varphi'(0)\cdot p_2/\varphi'(p_1) + o(p_2)$.

We already know that $p_1 = \mu_0(u)+o(x)$ and that $\mu_C(-u,x) = \mu_0(u)+o(x)$, so $\mu_C(u,x) - \mu_C(-u,C) = \varphi'(0)\cdot p_2/\varphi'(\mu_0(u)) + o(x)$. Substituting the known expression $p_2 = \mu_0'(\Delta) \cdot x \cdot \psi(\mu_0(u-\Delta)) + o(x)$ into this formula, and using the fact that $\mu_0(u)$ is an even function, we conclude that $\mu_C(u,x) - \mu_C(-u,x) = x \cdot \varphi'(0) \cdot \mu_0'(\Delta) \cdot \psi(\mu_0(\Delta-u))/\varphi'(\mu_0(u)) + o(x)$. Hence, the corresponding integral $N_1(x)$ can be represented as $x \cdot N_1'(0) + o(x)$, where

$$N_1'(0) = \mu_0'(\Delta) \cdot \int_0^\Delta \varphi'(0) \cdot \frac{u \cdot \psi(\mu_0(\Delta-u))}{\varphi'(\mu_0(u))}\,du.$$

We thus know the values of $N_1'(0)$, $N_2'(0)$, and $D(0)$. Hence, we can now compute the desirable value $u'(0) = N'(0)/D(0) = (N_1'(0) + N_2'(0))/D(0)$:

The resulting formulas.

$$|u'(0)| = |\mu_0'(\Delta)| \cdot \frac{A}{B},$$

where

$$B = 2\int_0^\Delta \mu_0(u)\,du,$$

and

$$A = \int_0^\Delta a(u)\,du,$$

with

$$a(u) = (u+\Delta) \cdot \psi(\mu_0(u)) + \varphi'(0) \cdot \frac{u \cdot \psi(\mu_0(\Delta-u))}{\varphi'(\mu_0(u))}.$$

These formulas are a modification of Sugeno integral. These formulas involve both the fuzzy logic operations (via the functions ψ and φ that describe these operations) *and* the normal arithmetic operations. Therefore, these formulas describe the desired modification of Sugeno integral.

Comment 1. The main *ideas* (but not the results) of such an estimation can be found in [4, 9]; the main difference between those papers and this new one being as follows:

- In [4, 9], we were looking for "and" and "or" operations that lead to the most stable and, correspondingly, the most smooth control.
- In this paper, we do not necessarily restrict ourselves to these two pairs of operations, because we realize that there are many other possible objectives of control. Instead, we analyze the degrees of stability and smoothness for an arbitrary choice of "and" and "or" operations.

Comment 2. It is desirable to relate our results to recently proposed modifications of Sugeno integral, e.g., to t-norm based fuzzy integral proposed by M. Grabisch, T. Murofushi, and M. Sugeno, and to Choquet-like modifications proposed by R. Mesiar, etc.

Acknowledgments.

This work was supported in part by NASA under cooperative agreement NCCW-0089, by NSF grants No. DUE-9750858, EEC-9322370, and CDA-9522207, and by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0518.

The authors are very grateful to Radko Mesiar and to the anonymous referees for valuable comments.

References

[1] M. Grabisch, H. T. Nguyen, and E. A. Walker, *Fundamentals of uncertainty calculi with applications to fuzzy inference*, Kluwer, Dordrecht, Netherlands, 1995.

[2] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory and applications*, Prentice Hall, Upper Saddle River, NJ, 1995.

[3] V. Kreinovich and H. T. Nguyen, Methodology of fuzzy control: an introduction, In [7].

[4] V. Kreinovich *et al.*, What non-linearity to choose? Mathematical foundations of fuzzy control, *Proceedings of the 1992 International Conference on Fuzzy Systems and Intelligent Control,* Louisville, KY, 1992, pp. 349–412.

[5] H. T. Nguyen and V. Kreinovich, A modification of Sugeno integral describes stability and smoothness of fuzzy control, *Proceedings of FUZZ-IEEE'98.*

[6] H. T. Nguyen, V. Kreinovich, and P. Wojciechowski, Strict Achimedean t-Norms and t-Conorms as Universal Approximators, *International Journal of Approximate Reasoning* (to appear).

[7] H. T. Nguyen and M. Sugeno (eds.), *Handbooks of Fuzzy Systems. Vol. VI: Fuzzy Modeling and Control*, Kluwer Academic, 1997 (to appear).

[8] H. T. Nguyen and E. A. Walker, *A first course in fuzzy logic*, CRC Press, Boca Raton, Florida, 1997.

[9] M. H. Smith and V. Kreinovich, Optimal strategy of switching reasoning methods in fuzzy control, In H. T. Nguyen *et al.* (eds.), *Theoretical aspects of fuzzy control*, J. Wiley, N.Y., 1995, pp. 117–146.

[10] M. Sugeno, *Theory of fuzzy integrals and its applications*, Doctoral Thesis, Tokyo Institute of Technology, 1974.

[11] M. Sugeno, Fuzzy measures and fuzzy integrals: a survey, In: Gupta, Saridis, and Gaines, eds., *Fuzzy Automata and Decision Processes*, 1977, pp. 89–102.

[12] L. Zadeh, Fuzzy sets, *Information and control*, 1965, Vol. 8, pp. 338–353.