Choosing t-norms and t-conorms for fuzzy controllers

Khurshid Ahmad
Department of Computer Science
Trinity College
Dublin, IRELAND
Khurshid.Ahmad@cs.tcd.ie

Andrea Mesiarová-Zemánková
Mathematical Institute
Slovak Academy of Sciences
Bratislava, SLOVAKIA
mesiar@mat.savba.sk

Abstract

Adaptive neuro-fuzzy systems rely critically on computing the output of a fuzzy system. The selection of t-norms for computing the 'firing strengths' of individual rules, and that of the dual t-conorms for computing the control output, relies on the conventional min/max t-norm/t-conorm dual combinations. This particular choice is form a host of combinations, including the product and the Łukasiewicz t-norms. In this paper we explore the differences in the outputs of two exemplar fuzzy systems - the Cartpole and the modelling of two dimensional sinc equation - using three different t-norm/t-conorm dual pairs. The implication of these results for the possible training of an ANFIS is discussed.

1. Introduction

The use of fuzzy sets in control systems has gained considerable theoretical interest (see, for instance, [3]) and the scope of applications of fuzzy control appears expanding as well ([2]). Fuzzy control systems are a generalization of classical control systems: it has been claimed that the fuzzy control method is a 'generalization of the human experience to use linguistic rules with vague predicates in order to formulate control actions' ([14]). The key to using linguistic rules, both in the premise and consequent part of the rules, attributed to Lotfi Zadeh, is a rather established use of t-norms and t-conorms for modelling the intersection and the union of fuzzy sets' [5]; A triangular norm (introduced in [8, 12]) is a binary operation $T : [0,1]^2 \to [0,1]$ such that for all $x, y, z \in [0,1]$ the following four axioms are satisfied:

(commutativity) $T(x,y) = T(y,x)$.
(associativity) $T(x,T(y,z)) = T(T(x,y),z)$.
(monotonicity) $T(x,y) \leq T(x,z)$ whenever $y \leq z$.
(boundary condition) $T(x,1) = x$.

The literature on triangular norms suggests that there are whole 'families' of t-norms, and their dual t-conorms - ranging from the familiar Gödelian minimum t-norm (and the associated maximum t-conorm) to the more general Archimedean t-norms. A well-grounded logical system should be based on a t-norm which is left-continuous: The continuity of the t-norm ensures stability of the fuzzy system. In implementing t-norms for computing firing strengths of individual rules, one can use t-conorms that may or may not be the dual of the t-norm. However, the choice of the dual t-conorm is to be preferred for the transparency of the implementation. Table 1 comprises the definitions of three widely cited triangular norms and t-conorms:

Triangular Norm	Gödelian minimum
Intersection Operator	$T_{\mathbf{M}}(a,b) = \min(a,b)$
Union Operator	$C_{\mathbf{M}}(a,b) = \max(a,b)$
Triangular Norm	Weber's Family, $\lambda \in [-1,\infty]$
Intersection Operator	$T_\lambda(a,b) = \max(\frac{a+b-1+\lambda ab}{1+\lambda}, 0)$
Union Operator	$C_\lambda(a,b) = \min(a+b-\frac{\lambda ab}{1+\lambda}, 1)$
Triangular Norm	Product ($\lambda \to \infty$)
Intersection Operator	$T_{\mathbf{P}}(a,b) = a \cdot b$
Union Operator	$C_{\mathbf{P}}(a,b) = a + b - a \cdot b$
Triangular Norm	Łukasiewicz ($\lambda = 0$)
Intersection Operator	$T_{\mathbf{L}}(a,b) = \max(0, a+b-1)$
Union Operator	$C_{\mathbf{L}}(a,b) = \min(a+b, 1)$

Table 1. Definitions of different t-norms and their dual t-conorms. ([6]).

It has been argued that the product t-norm is the strongest member of the Weber family and thus in some cases Yager family (which comprises from Drastic product t-norm to the minimum t-norm) can be better choice.

Advanced textbooks and monographs on neuro-fuzzy control systems ([10]), systems that are designed to learn the parameters of fuzzy membership functions and indeed fuzzy rules, suggest that 'instead of min another t-norm may be chosen, of course' ([10, p:207]): there is evidence that

Zadeh suggested the use of the min-max inference, minimum of the fuzzy membership values in evaluating the firing strength of a rule and the choice of maximum strength as indicated by the consequent part during composition, amongst other triangular norms. Such choice is rarely exercised although there are exceptions to the focus on min-max inference; in this context we will briefly review the work of Ciaramella et al [1] and of Pedrycz [11] later on in this paper.

2 Learning and additive generators

Given that there are parametric families of t-norms, and within each family one can choose a suitable t-norm for a control application by the right choice of parameter - in Table 1 above we showed how by choosing a the lambda parameter for the Weber family one can either have the product t-norm or Lukasiewicz t-norm. By choosing a 'correct' member one can reduce the computational complexity during the learning of the parameters of the membership functions of fuzzy variables. Note that each continuous Archimedean t-norm T (the t-norm T is called Archimedean if for each $(x, y) \in]0, 1[^2$ there is an $n \in \mathbb{N}$ with $x_T^{(n)} < y$, where $x_T^{(n)} = T(x, x_T^{(n-1)})$ and $x_T^{(1)} = x$) can be represented by means of a continuous additive generator [5, 7], i.e., a strictly decreasing continuous function $t : [0, 1] \to [0, \infty]$ with $t(1) = 0$ such that

$$T(x, y) = t^{(-1)}(t(x) + t(y)),$$

where the pseudo-inverse $t^{(-1)} : [0, \infty] \to [0, 1]$ in this special case is given by

$$t^{(-1)}(u) = t^{-1}(\min(u, t(0))).$$

A continuous Archimedean t-norm T is either strict, i.e., $T(x, y) = T(z, y)$ for $y > 0$ implies $x = z$ and then for an additive generator t of T we have $t(0) = \infty$, or nilpotent, i.e., for all $x \in [0, 1[$ there exists an $n \in \mathbb{N}$ such that $x_T^{(n)} = 0$ and then $t(0) < \infty$.

Additive generators for the product and the Łukasiewicz t-norms are $t_1(x) = -\ln(x)$ and $t_2(x) = 1 - x$, respectively. Weber t-norm with parameter $\lambda > 0$ has an additive generator given by $t(x) = 1-\log_{1+\lambda}(1+\lambda x)$. Note that the minimum t-norm does not have an additive generator (since it is not Archimedean).

Some families of t-norms are defined by applying the power transformation to the basic additive generator (additive generator of a family member which was chosen as a reference point). For example additive generators of t-norms from Yager family can be obtained by applying the power transformation to an additive generator of a Łukasiewicz t-norm (i.e., $t_2(x) = 1 - x$). In such a case computation of a parameter change (especially computation of a derivative in gradient descent method) is much easier.

3 Neuro-fuzzy Systems and triangular norms

We are focusing on the so-called neuro-fuzzy system, and especially adaptive neuro fuzzy systems as pioneered by Jang, Sun, and Mizutani [4]. Essentially, a fuzzy control system, and in particular a Takagi-Sugeno (T-S) [13] controller, is a network where the nodes act as union and intersection of linguistic variables, forming rules, and the links are weights related to the firing strengths of the rules and concomitantly to the parameters of the membership functions. The output layer of the network is the fuzzy systems equivalent of the defuzzification process - this is simplified significantly by either having the output as a constant value (zero-order T-S model) or output as a linear function of the inputs (first-order T-S model). An adaptive system is developed by providing a training set of patterns comprising inputs to and (desired) outputs from a system. The inputs are carried forward without any adjustment to the weights or the membership function parameters and in the defuzzification layer an error term is computed based on the computed and the desired output. This error computation is used to adjust the parameters of the consequent part of the rules using a least square estimation procedure. Such systems use neural-network based learning algorithms, especially the back-propagation algorithm, to propagate the error backwards in the network and in the process changes are made to the membership function parameters using, say, gradient descent methods. A study of the neuro fuzzy literature suggests that during the forward pass in an adaptive system, which is essentially the computation of the premise part of the rules, most authors use the Gödelian minimum t-norm.

Ciaramella et al have reported on the development of adaptive neuro-fuzzy systems. They have dealt specifically with such a system being based on a 'fuzzy relational "IF-THEN" reasoning scheme' [1, p:146] (or model), and the authors have defined the structure of the model using 'different t-norms and t-conorms'. The fuzzy-relational neural network model was implemented and was trained and tested to perform classification and approximation tasks - in both these tasks Ciaramella et al' system usually outperforms the widely used multi-layer perceptrons or neural networks based on the radial-basis function architectures. The authors cite t-norms (and t-conorms) listed in Table 1 together with the Yager: however, apart from a note (and a graph) showing 'classification [through] B^1 and B^2 fuzzy set estimation (sic)' for the well-known IRIS database, the different triangular norms, they simply conclude by suggesting that 'we obtain a 96% of correct classification on the training set.' in 'all [different norms] the cases' [1, p:156] . There is no overt mention of the triangular norms in the conclusion of this paper.

4 Cartpole and 'Sinc' simulation: two case studies

4.1 Cartpole problem

The first case study is from Cartpole problem frequently mentioned in the literature (see for example, [6]): Essentially, the problem is to balance a pole a smaller mass at its foot and a larger at its head. The amount of force required to balance a pole of length (l) during its movement is a function of the displacement of the pole by an angle θ and the angular velocity $\dot{\theta}$ and acceleration $\ddot{\theta}$ given as:

$$(M+m)\sin^2\theta\cdot l\cdot\ddot{\theta}+m\cdot l\cdot\sin\theta\cdot\cos\theta\cdot\theta^2 \\ -(M+m)\cdot g\cdot\sin\theta=-F\cos\theta,$$

where g is the acceleration due to gravity.

This equation can be solved numerically for demonstrating how feedback about the behavior in a previous time-step can be used to regulate the amount of force to keep the pole stable. More recently, it has been shown that a 'hierarchical reinforcement learning framework containing via-point representation' can be used to learn, and subsequently simulate the behavior of the Cartpole ([9, p:304]) is effective for learning from the definition of a trajectory and a novel reinforcement scheme. Fuzzy logic is soft-computing technique for 'informally' solving the cartpole problem through a set of linguistic variables embedded in rules - much like an expert may do ([6, p:163]). The set of 'reasonable' linguistic that can be stated about the behavior of Cartpole require the specification of term-sets of the variables angle, angular velocity and force - fuzzy partitions dividing the space of the variables into positive, negative and 'approximately zero' values, or more finer partitions comprising big positive (pb), medium positive (pm), and small positive (ps) values, big, medium and small negative values (nb, nm and ns respectively), together with a set of values that comprise the approximately zero (az) partition. These variables help to generate of 49 rules including rules like:

IF θ is ps and $\dot{\theta}$ is az **THEN** F is ps

Kruse et al have used 19 of the 49 rules in a simulation of the Cartpole problem and the rule base used by the authors is presented below in a matrix fashion: rows include the angular velocity term set and the columns the angle: each cell in the matrix comprises the consequent based on the two premises in the rows and columns:

	nb	nm	ns	az	ps	pm	pb
nb			ps	pb			
nm				pm			
ns	nm		ns	ps			
az	nb	nm	ns	az	ps	pm	pb
ps				ns	ps		pm
pm				nm			
pb				nb	ns		

Table 2: Rule base for the Cartpole problem (from [6, p:168])

The above rule-base matrix is quite sparse, especially for physically plausible combinations of the values of the variables in the premise there is usually one or two rules active at any given time. We have carried out three sets of simulations where at least two rules were active for a selection of plausible values of the angles and angular velocity resulting in the application of positive big, positive medium, positive small and approximately zero force:

	az	ps	pm
nb	pb		
nm	pm		
az	az	ps	pm
ps	ns	ps	

The 21 terms were mapped onto triangular membership functions covering: $\theta\in[-90,90]$, $\dot{\theta}\in[-45,45]$, and $F\in[-5,5]$.

We have used three sets of numbers ranging from a minimum to a maximum value carefully traversing over at most two rules - that is the values of the variable:

θ		$\dot{\theta}$	
Min	Max	Min	Max
2	22	-2	-10
2	20	-23	-33
24	44	0	9

This allows for the selection of two rules at a time, and the choice precluded most common-sensical of all rules - if there is no change (az) in the angle and in the angular velocity then there should be no change in the force (az).

The results of our computations for the difference in performance of the fuzzy control system when the Gödelian minimum was used versus the product t-norm (and its dual):

	ABs Pe	St dev	Max Error	Min Error	Max	Min
R1, R2	20.5%	22.8%	78.7%	0.1%	39.8%	-78.7%
R4, R5	4.3%	3.7%	13.8%	0.2%	0.2%	-13.8%
R1, R3	2.8%	3.5%	13.3%	0.0%	6.5%	-13.3%

where rule

R1	is	ps & az	$\rightarrow$	ps,
R2	is	ps & ps	$\rightarrow$	ps,
R3	is	pm & az	$\rightarrow$	pm,
R4	is	az & nm	$\rightarrow$	pm,
R5	is	az & nb	$\rightarrow$	pb.

Note that the structure of this fuzzy model induces that the biggest error (difference) can be expected in the area where the output of the system is near to zero. The absolute percentage error, and other corresponding metrices like standard deviation of the differences between the output of two systems, shows a major effect for one set of rules where the values chosen above shows significant variation over one or more partitions - positive small has an overlap with positive medium and approximately zero.

4.2 Sinc function

In the second study we have focused to the model introduced by Jang, Sun, and Mizutani [4], which is an approximation of a two dimensional sinc equation. The partition in this model is the same for both input variables - for each we have used four bell-shaped fuzzy sets with parameters

a	b	c
-10	2	4
-3	2	3
3	2	3
10	2	4

The results of our computations for the difference in performance of the fuzzy system when the Gödelian minimum was used versus the product t-norm (and its dual):

	Min			Prod		
	Abs. Error	Diff	Perc. Error	Abs. Error	Diff	Perc. Error
Stdev	0.33	0.37	133.68	0.35	0.39	140
MSE		0.03			0.04	
RMSE		0.18			0.19	

Recall that the crucial task in this computation is the choice of the consequence parameters and this is the task on which we have focused in further development of this model. Note that the choice can by e.g. done by linearization of a sinc function in the points which are cores of input fuzzy sets.

5 Afterword and Future Work

It is perhaps true that despite the differences in the properties of the triangular norms of a range of families, the differences between the minimum value of a conjunction of membership functions and a product of the same will be minimal for vanishing membership value of any of the linguistic variables ($T(0,x) = T(x,0) = 0$) or the case when the membership values of any of the variables peaks to unity ($T(1,x) = x$). So, there is no difference between the triangular norms at these idiosyncratic values of the variables, then the inferencing performed using the norms will essentially be the same. We show this to be case for two particular exemplars later on. But what is the import of the statement that even when different norms and t-conorms are used for inferencing (and composition), the results, for example, control values, are essentially the same? One possible answer is that when we are very certain about the 'belongingness' or otherwise of a variable to fuzzy set, what we have is a classical exclusive belongingness relationships: here the whole paraphernalia of fuzzy systems (Fuzzification $\longrightarrow$ Inference $\longrightarrow$ Composition $\longrightarrow$ [Defuzzification]) is extraneous and we need not worry about fuzziness or fuzzy systems here. But there are key differences in the output of fuzzy control systems when the variables do not have the idiosyncratic values. This we have demonstrated by the two case studies.

Our work continues on building a system that will be a full implementation of Jang, Sun and Mitzuani's ANFIS algorithm as our attempt to delineate the effects of the choice of a typical triangular norm and its dual. Note that our consideration will further be extend from Archimedean t-norms also to ordinal sums of t-norms (compare [11]).

References

[1] Ciaramella, A., Tagliaferri, R., Pedrycz, W., and di Nola, A. (2006). 'Fuzzy relational network'. International Journal of Approximate Reasoning. Vol 41, pp 146-163.

[2] Feng, Gang (2006). A Survey on Analysis and Design of Model-Based Fuzzy Control Systems. IEEE Transactions on Fuzzy Systems, Vol 14 (No.5), pp 676 - 697

[3] Gerla, Giangiacomo ((2005) Fuzzy logic programming and fuzzy control, Studia Logica, Vol 79 (No. 2), pp. 231-254.

[4] Jang, Roger J-S., Sun, Cheun-Tsai., and Mizutani, Eiji. (1997). Neuro-fuzzy and Soft Computing. Upper Saddle River (New Jersey): Prentice-Hall, Inc.

[5] Klement, Erich, P., Mesiar, Radko., and Pap Endre. (2000). Triangular Norms. Dodrecht: Kluwer Academic Publishers.

[6] Kruse, Rudolf., Gebhardt, Jrg., and Klawonn, Frank. (1994). Foundations of Fuzzy Systems. Chichester: John Wiley & Sons.

[7] C. M. Ling (1965). Representation of associative functions. *Publ. Math. Debrecen*, **12**, pp. 189–212.

[8] K. Menger (1942). Statistical metrics. *Proc. Nat. Acad. Sci. U.S.A.* **8**, pp. 535–537.

[9] Miyamoto, Hiroyuki, Jun Morimoto, Kenji Doya, Mitsuo Kawato. (2004). 'Reinforcement learning with via-point representation'. Neural Networks Vol 17 , 299-305

[10] Nauck, Detlefand., Klawonn, Frank., and Kruse, Rudolf (1998). Foundations of Neuro-fuzzy Systems. Chichester: John Wiley & Sons.

[11] Pedrycz, Witold (2006). Logic-Based Fuzzy Neurocomputing With Unineurons, IEEE Transactions on Fuzzy Systems, Vol 14 (No.6), pp 860-873.

[12] B. Schweizer and A. Sklar (1960). Statistical metric spaces. *Pacific J. Math.*, **10**, pp. 313–334.

[13] Takagi, T., & Sugeno, M. (1985). 'Fuzzy Identification of Systems and its Applications to Modeling and Control'. IEEE Transactions on Systems, Man and Cybernetics. Volume No. SMC-15 (No.1) pp 116-132.

[14] Yager, R R & Filev, D P. (1994) Essentials of Fuzzy Modeling and Control. Chichester: John Wiley & Sons Ltd.