Budapest Tech
Hungarian Fuzzy Association

5th International Symposium of Hungarian Researchers

on

COMPUTATIONAL INTELLIGENCE

PROCEEDINGS

Magyar Kutatók 5. Nemzetközi Szimpóziuma

2004
NOVEMBER 11-12
BUDAPEST

Budapesti Műszaki Főiskola
Magyar Fuzzy Társaság

The Sigmoidal Approximation of Łukasiewicz Operators

József Dombi, Zsolt Gera
University of Szeged, Institute of Informatics
e-mail: {dombi | gera}@inf.u-szeged.hu

Abstract: In this paper we propose an approximation of Łukasiewicz operators by means of sigmoid functions. Łukasiewicz operators play an important role in fuzzy logic. They are widely used due to their good theoretical properties, i.e. the residual and material implications coincide, the law of excluded middle and the law of non-contradiction both hold. Besides these good theoretical properties this operator family does not have a continuous gradient. Its approximation is simple and continuously differentiable.
Keywords: approximation, sigmoid function, Łukasiewicz operators

1 Introduction

The Łukasiewicz operator class (see e.g. [1], [2], [3]) is commonly used for various purposes. In this well known operator family the cut function (denoted by $[\cdot]$) plays an important role. We can get the cut function from x by taking the maximum of 0 and x and then taking the minimum of the result and 1.

Definition 1.1. *Let the **cut function** be*

$$[x] = \min(\max(0, x), 1) = \begin{cases} 0, & \text{if } x \le 0 \\ x, & \text{if } 0 < x < 1 \\ 1, & \text{if } 1 \le x \end{cases} \tag{1}$$

*Let the **generalized cut function** be*

$$[x]_{a,b} = [(x-a)/(b-a)] = \begin{cases} 0, & \text{if } x \le a \\ \frac{x-a}{b-a}, & \text{if } a < x < b \\ 1, & \text{if } b \le x \end{cases} \tag{2}$$

where $a, b \in \mathbb{R}$ and $a < b$.

All nilpotent operators are constructed using the cut function. The formulas of the nilpotent conjunction, disjunction, implication and negation are the following:

$$\begin{aligned} c(x,y) &= [x+y-1] \\ d(x,y) &= [x+y] \\ i(x,y) &= [1-x+y] \\ n(x) &= 1-x \end{aligned} \tag{3}$$

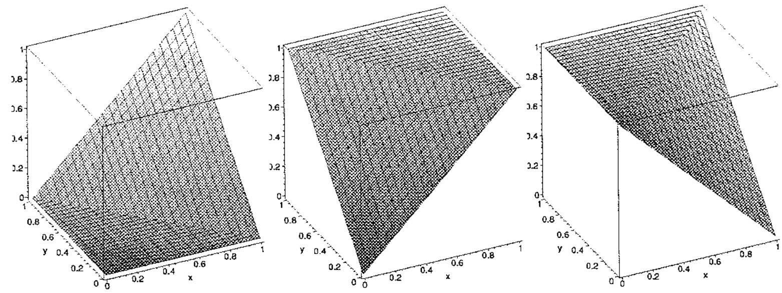

Figure 1: The truth tables of the nilpotent conjunction, disjunction and implication

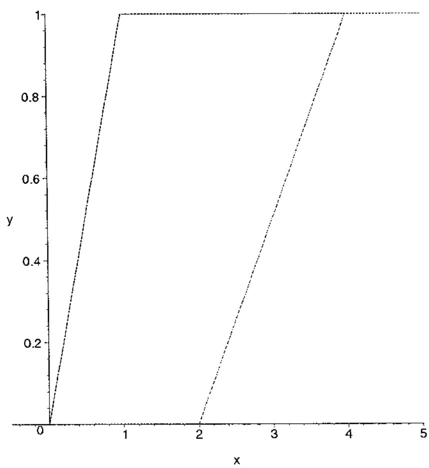

Figure 2: Two generalized cut functions.

The truth tables of the former three can be seen on fig. 1.

The Łukasiewicz operator family used above has good theoretical properties. These are: the law of non-contradiction (that is the conjunction of a variable and its negation is always zero) and the law of excluded middle (that is the disjunction of a variable and its negation is always one) both hold, and the residual and material implications coincide. These properties make these operators to be widely used in fuzzy logic and to be the closest one to classic Boolean logic. Besides these good theoretical properties this operator family does not have a continuous gradient. So for example gradient based optimization techniques are impossible with Łukasiewicz operators. The root of this problem is the shape of the cut function itself.

2 Approximation of the Cut Function

A solution to above mentioned problem is a continuously differentiable approximation of the cut function, which can be seen on fig. 3. In this section we'll construct such an approximating function by means of sigmoid functions. The reason for choosing the sigmoid function was that this function has a very important role in many areas. It is frequently used in artificial neural networks, optimization methods, economical and biological models.

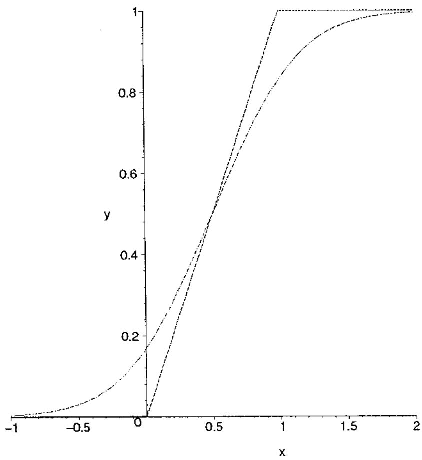

Figure 3: The cut function and its approximation

2.1 The Sigmoid Function

The sigmoid function (see fig. 4) is defined as

$$\sigma_d^{(\beta)}(x) = \frac{1}{1 + e^{-\beta(x-d)}} \tag{4}$$

where the lower index d is omitted if 0.

Let us examine some of its properties which will be useful later:

- its derivative can be expressed by itself (see fig. 5):

$$\frac{\partial \sigma_d^{(\beta)}(x)}{\partial x} = \beta \sigma_d^{(\beta)}(x) \left(1 - \sigma_d^{(\beta)}(x)\right) \tag{5}$$

- its integral has the following form:

$$\int \sigma_d^{(\beta)}(x)\, dx = -\frac{1}{\beta} \ln \left(\sigma_d^{(-\beta)}(x)\right) \tag{6}$$

Because the sigmoid function is asymptotically 1 as x tends to infinity, the integral of the sigmoid function is asymptotically x (see fig. 6).

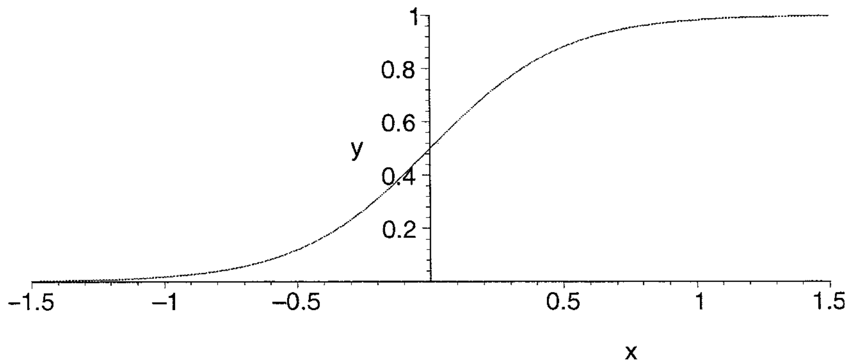

Figure 4: The sigmoid function, with parameters $d = 0$ and $\beta = 4$

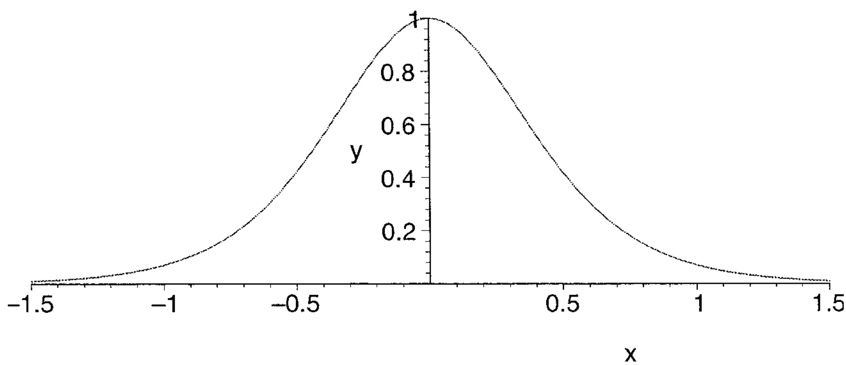

Figure 5: The first derivative of the sigmoid function

2.2 The Squashing Function on $[a, b]$

In order to get an approximation of the generalized cut function, let us integrate the difference of two sigmoid functions, which are translated by a and b ($a < b$), respectively.

$$
\begin{aligned}
&\frac{1}{b-a}\int \sigma_a^{(\beta)}(x) - \sigma_b^{(\beta)}(x)\,dx = \\
&= \frac{1}{b-a}\left(\int \sigma_a^{(\beta)}(x)\,dx - \int \sigma_b^{(\beta)}(x)\,dx\right) = \\
&= \frac{1}{b-a}\left(-\frac{1}{\beta}\ln\left(\sigma_a^{(-\beta)}(x)\right) + \frac{1}{\beta}\ln\left(\sigma_b^{(-\beta)}(x)\right)\right)
\end{aligned}
$$

After simplification we get the squashing function on the interval $[a, b]$:

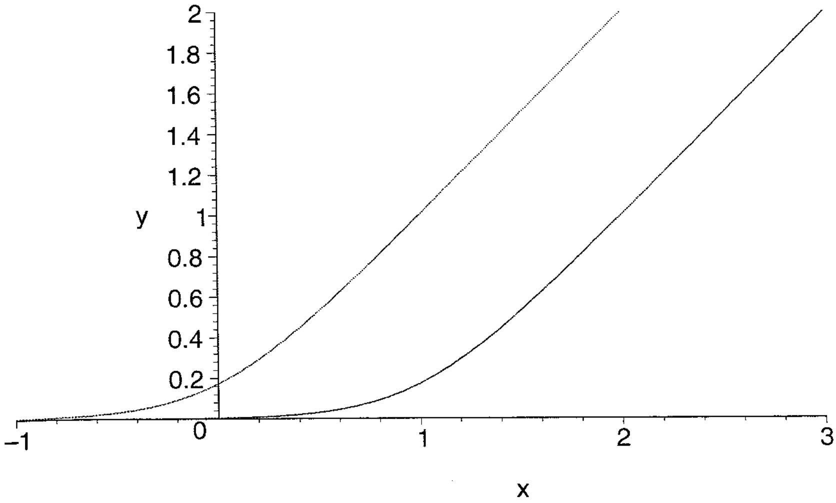

Figure 6: The integral of the sigmoid function, one is shifted by 1

Definition 2.1. *Let the **interval squashing function** on $[a, b]$ be*

$$S_{a,b}^{(\beta)}(x) = \frac{1}{b-a} \ln \left(\frac{\sigma_b^{(-\beta)}(x)}{\sigma_a^{(-\beta)}(x)} \right)^{1/\beta} = \frac{1}{b-a} \ln \left(\frac{1+e^{\beta(x-a)}}{1+e^{\beta(x-b)}} \right)^{1/\beta} . \quad (7)$$

The parameters a and b affect the placement of the interval squashing function, while the β parameter drives the precision of the approximation. We need to prove that $S_{a,b}^{(\beta)}(x)$ is really an approximation of the generalized cut function.

Theorem 2.2. *Let $a, b \in \mathbb{R}$, $a < b$ and $\beta \in \mathbb{R}^+$. Then*

$$\lim_{\beta \to \infty} S_{a,b}^{(\beta)}(x) = [x]_{a,b} \quad (8)$$

and $S_{a,b}^{(\beta)}(x)$ is continuous in x, a, b and β.

Proof. It is easy to see the continuity, because $S_{a,b}^{(\beta)}(x)$ is a simple composition of continuous functions and because the sigmoid function has a range of $[0, 1]$ the quotient is always positive.

In proving the limit we separate three cases, depending on the relation between a, b and x.

- Case 1 ($x < a < b$): Since $\beta(x - a) < 0$, so $e^{\beta(x-a)} \longrightarrow 0$ and similarly $e^{\beta(x-b)} \longrightarrow 0$. Hence the quotient converges to 1 if $\beta \longrightarrow \infty$, and the logarithm of one is zero.

- Case 2 ($a \leq x \leq b$):

$$\frac{1}{b-a}\ln\left(\lim_{\beta\to\infty}\left(\frac{1+e^{\beta(x-a)}}{1+e^{\beta(x-b)}}\right)^{1/\beta}\right) =$$

$$= \frac{1}{b-a}\ln\left(\lim_{\beta\to\infty}\left(\frac{e^{\beta(x-a)}\left(e^{-\beta(x-a)}+1\right)}{\left(1+e^{\beta(x-b)}\right)}\right)^{1/\beta}\right) =$$

$$= \frac{1}{b-a}\ln\left(\lim_{\beta\to\infty}\frac{e^{x-a}\left(e^{-\beta(x-a)}+1\right)^{1/\beta}}{\left(1+e^{\beta(x-b)}\right)^{1/\beta}}\right) =$$

$$= \frac{1}{b-a}\ln\left(e^{x-a}\lim_{\beta\to\infty}\frac{\left(e^{-\beta(x-a)}+1\right)^{1/\beta}}{\left(1+e^{\beta(x-b)}\right)^{1/\beta}}\right)$$

We transform the nominator so that we can take the e^{x-a} out of the limes. In the nominator $e^{-\beta(x-a)}$ remained which converges to 0 as well as $e^{\beta(x-b)}$ in the denominator so the quotient converges to 1 if $\beta \to \infty$. So as the result, the limit of the interval squashing function is $(x-a)/(b-a)$, which by definition equals to the generalized cut function in this case.

- Case 3 ($a < b < x$):

$$\frac{1}{b-a}\ln\left(\lim_{\beta\to\infty}\left(\frac{1+e^{\beta(x-a)}}{1+e^{\beta(x-b)}}\right)^{1/\beta}\right) =$$

$$= \frac{1}{b-a}\ln\left(\lim_{\beta\to\infty}\left(\frac{e^{\beta(x-a)}\left(e^{-\beta(x-a)}+1\right)}{e^{\beta(x-b)}\left(e^{-\beta(x-b)}+1\right)}\right)^{1/\beta}\right) =$$

$$= \frac{1}{b-a}\ln\left(\lim_{\beta\to\infty}\frac{e^{x-a}\left(e^{-\beta(x-a)}+1\right)^{1/\beta}}{e^{x-b}\left(e^{-\beta(x-b)}+1\right)^{1/\beta}}\right) =$$

$$= \frac{1}{b-a}\ln\left(\frac{e^{x-a}}{e^{x-b}}\lim_{\beta\to\infty}\frac{\left(e^{-\beta(x-a)}+1\right)^{1/\beta}}{\left(e^{-\beta(x-b)}+1\right)^{1/\beta}}\right)$$

We do the same transformations as in the previous case but we take e^{x-b} from the denominator, too. After these transformations the remaining quotient converges to 1, so

$$\lim_{\beta\to\infty} S_{a,b}^{(\beta)}(x) = \frac{1}{b-a}\ln\left(\frac{e^{x-a}}{e^{x-b}}\right) = \frac{1}{b-a}\ln\left(e^{x-a-(x-b)}\right) =$$

$$= \frac{1}{b-a}\ln\left(e^{b-a}\right) = \frac{b-a}{b-a} = 1.$$

■

On fig. 7 the interval squashing function can be seen with various β parameters. The following proposition states some properties of the interval squashing function.

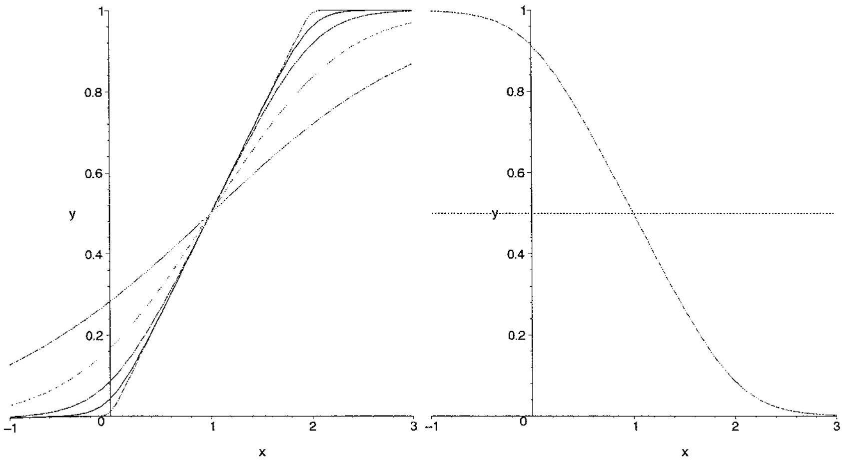

Figure 7: On the left image: the interval squashing function with an increasing β parameter ($a = 0$ and $b = 2$). On the right image: the interval squashing function with a zero and a negative β parameter

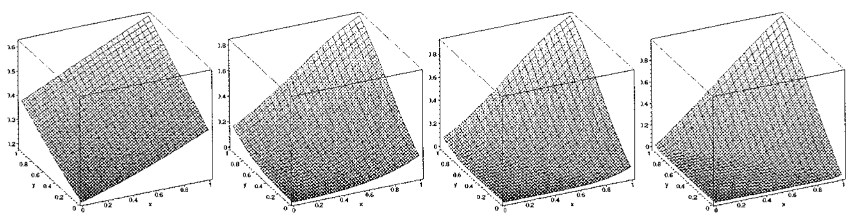

Figure 8: The approximation of the nilpotent conjunction with β values 1,4,8 and 32

Proposition 2.3.

$$\lim_{\beta \to 0} S_{a,b}^{(\beta)}(x) = 1/2 \tag{9}$$

$$S_{a,b}^{(-\beta)}(x) = 1 - S_{a,b}^{(\beta)} \tag{10}$$

As an another example, the nilpotent conjunction is approximated with the interval squashing function on fig. 8.

For further use, let us introduce an another form of the interval squashing function's formula. Instead of using parameters a and b which were the "bounds" on the x axis, from now on we'll use a and δ, where a gives the center of the squashing function and where δ gives its steepness. Together with the new formula we introduce its pliant notation.

Definition 2.4. *Let the* ***squashing function*** *be*

$$\langle a <_\delta x \rangle_\beta = S_{a,\delta}^{(\beta)}(x) = \frac{1}{2\delta} \ln \left(\frac{\sigma_{a+\delta}^{(-\beta)}(x)}{\sigma_{a-\delta}^{(-\beta)}(x)} \right)^{1/\beta}, \tag{11}$$

where $a \in \mathbb{R}$ *and* $\delta \in \mathbb{R}^+$.

If the a *and* δ *parameters are both* $1/2$ *we will use the following pliant notation for simplicity:*

$$\langle x \rangle_\beta = S_{\frac{1}{2},\frac{1}{2}}^{(\beta)}(x), \tag{12}$$

which is the approximation of the cut function.

The inequality relation in the pliant notation refers to the fact that the squashing function can be interpreted as the truthness of the relation $a < x$ with decision level $1/2$, according to a fuzziness parameter δ and an approximation parameter β (see fig. 9).

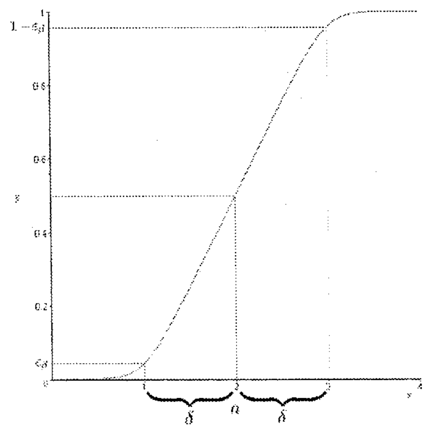

Figure 9: The meaning of $\langle a <_\delta x \rangle_\beta$

The derivatives of the squashing function can be expressed by itself and sigmoid functions:

$$\frac{\partial S_{a,\delta}^{(\beta)}(x)}{\partial x} = \frac{1}{2\delta} \left(\sigma_{a-\delta}^{(\beta)}(x) - \sigma_{a+\delta}^{(\beta)}(x) \right) \tag{13}$$

$$\frac{\partial S_{a,\delta}^{(\beta)}(x)}{\partial a} = \frac{1}{2\delta} \left(\sigma_{a+\delta}^{(\beta)}(x) - \sigma_{a-\delta}^{(\beta)}(x) \right) \tag{14}$$

$$\frac{\partial S_{a,\delta}^{(\beta)}(x)}{\partial \delta} = \frac{1}{2\delta} \left(\sigma_{a+\delta}^{(\beta)}(x) + \sigma_{a-\delta}^{(\beta)}(x) \right) - \frac{1}{\delta} S_{a,\delta}^{(\beta)}(x) \tag{15}$$

2.3 The Error of the Approximation

The squashing function approximates the cut function with an error. This error can be defined in many ways. We have chosen the following definition.

Definition 2.5. *Let the approximation error of the squashing function be*

$$\varepsilon_\beta = \langle 0 <_\delta (-\delta) \rangle = \frac{1}{2\delta} \ln \left(\frac{\sigma_\delta^{(-\beta)}(-\delta)}{\sigma_{-\delta}^{(-\beta)}(-\delta)} \right)^{1/\beta} \tag{16}$$

where $\beta > 0$.

Because of the symmetry of the squashing function $\varepsilon_\beta = 1 - \langle 0 <_\delta \delta \rangle$, see fig. 9.

The purpose of measuring the approximation error is the following inverse problem: we want to get the corresponding β parameter for a desired ε_β error. We state the following lemma on the relationship between ε_β and β.

Lemma 2.6. *Let us fix the value of* δ. *The following holds for* ε_β.

$$\varepsilon_\beta < c \cdot \frac{1}{\beta}, \tag{17}$$

where $c = \frac{\ln 2}{2\delta}$ *is constant.*

Proof.

$$\varepsilon_\beta = \frac{1}{2\delta\beta} \ln \left(\frac{1 + e^{\beta(-\delta+\delta)}}{1 + e^{\beta(-\delta-\delta)}} \right) = \frac{1}{2\delta\beta} \ln \left(\frac{2}{1 + e^{-2\delta\beta}} \right) =$$
$$= \frac{\ln 2}{2\delta\beta} - \frac{\ln(1 + e^{-2\delta\beta})}{2\delta\beta} < c \cdot \frac{1}{\beta}$$

■

So the error of the approximation can be upper bounded by $c \cdot \frac{1}{\beta}$, which means that by increasing parameter β, the error decreases by the same order of magnitude.

Conclusion

In this paper we have reviewed the cut function, which is the basis of the well known Łukasiewicz operator class. This cut function is piecewise linear, it can not be continuously differentiated. We have created an approximation of the cut function (the squashing function) by means of sigmoid functions with good analytical properties, for example fast convergence and easy calculation.

References

[1] R. Ackermann. *An Introduction to Many-Valued Logics.* Dover, New York, 1967.

[2] P. Hájek. *Metamathematics of Fuzzy Logic.* Kluwer, 1998.

[3] D. Mundici R. Cignoli, I. M. L. D'Ottaviano. Algebraic foundations of many-valued reasoning. *Trends in Logic*, 7, 2000.