

Maximum Tolerance and Maximum Greatest Tolerance

J. Freixas

EPSEM-UPC. E-08240 Manresa, Spain
Dpt. Matemàtica Aplicada 3
E-mail: josep.freixas@upc.edu

X. Molinero

EPSEM-UPC. E-08240 Manresa, Spain
Dpt. Llenguatges i Sistemes Informàtics
E-mail: molinero@lsi.upc.edu

September 16, 2009

Abstract

An important consideration when applying neural networks is the sensitivity to weights and threshold in strict separating systems representing a linearly separable function. Two parameters have been introduced to measure the relative errors in weights and threshold of strict separating systems: the tolerance and the greatest tolerance. Given an arbitrary separating system we study which is the equivalent separating system that provides maximum tolerance or/and maximum greatest tolerance.

Keywords: Neural nets, circuit-switching networks, deterministic and structural pattern recognition.

1 Introduction

Research in threshold logic synthesis became an area of great interest and was done mostly in the 1950s and 1960s. Approximation methods are used in (Winder, 1962), (Dertouzos, 1965) and (Hu, 1965) to determine the input weights and threshold of a threshold function. Linear programming and tabulation methods have been used in (Muroga, 1971) and (Hammer et al., 1981) to determine if a function is threshold or not. However, Parberry (Parberry, 1994) clearly quotes that “it would be unreasonable to expect that natural or artificial neurons are able to realize every linear threshold function”.

Recent research in Capacitive Threshold Logic (Sang-Hoon and Lee, 1995), (Ozdemir et al., 1996) and (Beiu et al., 2003) has revived interest in this area, and it has re-introduced some of the problems that have yet to be solved. One of the main issues of threshold logic is the application of neural networks to the problem of realizing Boolean functions, the linear separability problem has been dealt with, among others, in (Yao and Ostapko, 1968), (Roychowdhury et al., 1994), (Siu et al.,

1995), (Picton, 1991), (Elizondo, 2004) and (Elizondo, 2006). Neural networks are usually designed to have the ability to learn and generalize.

A neural network is capable of setting and input–output mapping by adjusting its threshold and weights. How can network designers predict the effect of threshold and weight perturbations on neural networks’ output, which are unavoidable because of limited precision of digital and analog hardware? What is then a tight bound for weights and threshold to prevent output deviation? Which is the sharpest bound one may consider for weights and threshold to maintain the linearly separable function unchangeable? Is there any guideline for designing a more robust and safer neural network? These kind of problems were deeply studied in the sixties (Hu, 1960), (Elgot, 1961), (Myhill and Kautz, 1961) and (Winder, 1962), until (Hu, 1965) proposed as a solution to the main problem, a number (defined for each strict separating system) which he called, the *tolerance*. Recently (Freixas and Molinero, 2008a) proposed a new bound which improves the tolerance, *the greatest tolerance* (*G*-tolerance), and proved that the *G*-tolerance is the greatest bound one may consider.

In this work we consider two parameters for an arbitrary linearly separable switching function: the number of variables n and the number of types¹ of distinguished variables k . For each pair (n, k) we are interested in determining the maximum achievable value for the tolerance and for the *G*-tolerance. Moreover, we demonstrate that taking strict separating systems with positive integer (natural) weights are enough to our purpose.

2 Preliminaries

Let Q be the set $\{0, 1\}$. For any given positive integer n , consider the cartesian power product $Q^n = Q \times \dots \times Q$. Thus, the elements of Q^n are the 2^n ordered n -tuples (x_1, \dots, x_n) , with variables $x_i \in \{0, 1\}$ for all $i = 1, \dots, n$. By a *switching function* of n variables, we mean a function $f : Q^n \rightarrow Q$ from the n -cube Q^n into Q . If f is not surjective then either $\{x \in Q^n : f(x) = 0\} = \emptyset$ or $\{x \in Q^n : f(x) = 1\} = \emptyset$ that is, f is a constant function. The two constant functions are called the two *trivial* switching functions. A switching function is *monotonic* if: (i) f is not decreasing in each argument; and (ii) $f(\bar{0}) = 0$ and $f(\bar{1}) = 1$. The two constant Boolean functions are usually considered to be monotonic, but here the restriction (ii) relegates them to be non-monotonic. In a monotonic switching function a variable i is *relevant* if there exists at least $x, y \in Q^n$, $x \geq y$ with $f(x) = 1$, $f(y) = 0$, $x_i \neq y_i$ and $y_j = x_j$ for $j \neq i$. Note that *irrelevant* (that is, not relevant) variables do not add any value to the outcome f , i.e., if i is irrelevant then $f(x) = f(y)$ whenever $x_i \neq y_i$ and $y_j = x_j$ for $j \neq i$.

A switching function $f : Q^n \rightarrow Q$ is a *linearly separable function* or *threshold function* if it admits a system of $n + 1$ real numbers T, w_1, \dots, w_n , denoted by $[T; w_1, \dots, w_n]$ such that for each arbitrary point $x = (x_1, \dots, x_n)$ in the n -cube Q^n

¹A type is a set of variables which are equivalent among them.

we have

$$\begin{aligned} w(x) &\geq T, & \text{if } f(x) = 1, \\ w(x) &< T, & \text{if } f(x) = 0; \end{aligned}$$

being $w(x) = \sum_{i=1}^n w_i x_i$.

The n real numbers w_1, \dots, w_n in this system are called the *weights*, and the first real number T is referred to as the *threshold*. By the finiteness of Q^n it is always possible to modify the threshold in such a way that the previous definition could be rewritten using strict inequalities. In this case, the system is called a *strict separating system* for the linear separable function f . Thus, from now on we will just consider strict separating systems.

Note that a switching function has only two possible values, true or false. True or false can also be referred, as we do here, to as 1 or 0 often used in computer programming, on or off as seen with computer hardware circuits, firing or resting used to describe the state of an artificial neuron, functioning state or failing state as seen with reliability, and voting in favor (“yea”) or against (“nay”) in binary decision-making mechanisms. If one draws the graph of a linearly separable switching function of two variables in a square with the elements of Q^2 as vertices, it takes only one straight line to separate the true outputs from the false outputs. In general, for n variables it is required an $(n - 1)$ -hyperplane to separate the true outputs from the false outputs.

It is important to point out that both, the tolerance and the G -tolerance, we are going to introduce next, are defined for each strict separating system and, thus they depend on the threshold and the set of weights chosen to implement the linearly separable switching function.

2.1 The Tolerance

In this part we are going to define the tolerance introduced by (Hu, 1960). Let A denote the maximum of the function $w(x)$ for all x such that $f(x) = 0$,

$$A = \max_{x: f(x)=0} w(x)$$

and let B denote the minimum of the function $w(x)$ for all x such that $f(x) = 1$,

$$B = \min_{x: f(x)=1} w(x).$$

The two trivial switching functions are linearly separable; if $f \equiv 0$, we set $A = -\infty$; if $f \equiv 1$, we set $B = \infty$. Then we have $A < T < B$. Let m denote the smallest of the two positive numbers $T - A$ and $B - T$. On the other hand, let $M = |T| + \sum_{i=1}^n |w_i|$. Then, for each point $x \in Q^n$ we have $|T| + \sum_{i=1}^n |w_i| x_i \leq M$. Let $\lambda_1, \dots, \lambda_n$ and Λ be $n + 1$ arbitrary real numbers and let

$$\begin{aligned} w'_i &= (1 + \lambda_i) w_i, & \text{for all } i = 1, \dots, n \\ T' &= (1 + \Lambda) T. \end{aligned}$$

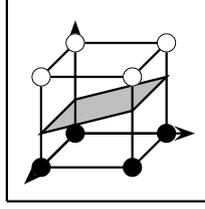


Figure 1: Separating hyperplane for a threshold function defined in Q^3 .

Then, the real numbers $\lambda_1, \dots, \lambda_n$ and Λ represent the relative errors if we use the numbers w'_1, \dots, w'_n and T' instead of the original non-null numbers w_1, \dots, w_n and T as weights and threshold. In other words, if we initially start with $[T; w_1, \dots, w_n]$ which is transformed into $[T'; w'_1, \dots, w'_n]$ then $\lambda_i = \frac{w'_i - w_i}{w_i}$ for all $i = 1, \dots, n$ and $\Lambda = \frac{T' - T}{T}$ where it is required w_i and T to be different from zero. That is to say, these numbers are the *relative errors* in weights and threshold, that is why we use the term “error” in the title of this paper.

Theorem 2.1 (Hu, 1965)

Let $f : Q^n \rightarrow Q$ be an arbitrary linearly separable switching function of n variables, let $[T; w_1, \dots, w_n]$ be a given strict separating system² for f and let $\tau[T; w_1, \dots, w_n] := \frac{m}{M}$ be the tolerance for $[T; w_1, \dots, w_n]$. If $|\lambda_i| < \tau$ for each $i = 1, \dots, n$ and if $|\Lambda| < \tau$ then $[T'; w'_1, \dots, w'_n]$ is a strict separating system for the given linearly separable switching function f .

Hu called this positive number the *tolerance* of the separating system and denoted it $\tau[T; w_1, \dots, w_n]$.

Note that the tolerance is well-defined, in fact, as $w(\bar{0}) = 0$ for any strict separating system, it occurs that $T \neq 0$ and so $M \neq 0$.

2.2 The Greatest Tolerance

(Freixas and Molinero, 2008a) improves the *Hu's tolerance*. They find the greatest positive real number δ such that if

$$|\Lambda| < \delta \quad \text{and} \quad |\lambda_i| < \delta, \quad \text{for all } i = 1, \dots, n$$

then $[T'; w'_1, \dots, w'_n]$ is equivalent to $[T; w_1, \dots, w_n]$, i.e., both strict separating systems still represent the same linearly separable switching function.

Definition 2.2 (Freixas and Molinero, 2008a)

² $\tau[T; w_1, \dots, w_n]$ is denoted by τ if there is no confusion about the strict linearly separating system.

Given a strict separating system $[T; w_1, \dots, w_n]$, for each $x \in Q^n$ let $a(x) = |w(x) - T|$, $b(x) = |T| + \sum_{i=1}^n |w_i| x_i$ and $\chi[T; w_1, \dots, w_n] = \min_{x \in Q^n} \frac{a(x)}{b(x)}$. This number³ is called the greatest tolerance (briefly, *G-tolerance*) of $[T; w_1, \dots, w_n]$. Note that the *G-tolerance* of the strict separating system $[T; w_1, \dots, w_n]$ depends on the chosen threshold and weights. Because of the finiteness of Q^n , χ is attained for, at least one $x \in Q^n$, let $x^0 \in Q^n$ be one of the points attaining the *G-tolerance*.

Theorem 2.3 (Freixas and Molinero, 2008a)

Let $f : Q^n \rightarrow Q$ be an arbitrarily linearly separable switching function of n variables and let $[T; w_1, \dots, w_n]$ be a given strict separating system for f . If $|\lambda_i| < \chi$ for each $i = 1, \dots, n$ and if $|\Lambda| < \chi$ then:

- (i) $[T'; w'_1, \dots, w'_n]$ is a strict separating system for the given linearly separable switching function f .
- (ii) χ is the greatest upper bound for the constants $\lambda_1, \dots, \lambda_n, \Lambda$.

3 Maximum Tolerance and Maximum *G-Tolerance*

Let f be an arbitrary monotonic⁴ linearly separable switching function of n variables, now we are looking for a strict linearly⁵ separating system $[T; w_1, \dots, w_n]$ representing f with *maximum achievable tolerance* and *maximum achievable G-tolerance*.

We start by relating the tolerance (the *G-tolerance*) of an arbitrary strict separating system with the tolerance (*G-tolerance*) of a strict separating system with natural weights.

Theorem 3.1

If $[T; w_1, \dots, w_n] \in \mathbb{R} \times \mathbb{R}^n$ is a real strict separating system, then there exists an equivalent strict separating natural system, i.e., a system with natural weights w'_1, \dots, w'_n , with the same tolerance and the same *G-tolerance*.

Note that, even though the weights are the same natural numbers for both systems (the system applied to the tolerance and the system applied to the *G-tolerance*), the threshold can be a different real number for each system.

An obvious corollary for maximums arises:

Corollary 3.2

If $[T; w_1, \dots, w_n] \in \mathbb{R} \times \mathbb{R}^n$ is a real strict separating system with maximum tolerance (maximum *G-tolerance*), then there exists an equivalent strict separating natural system with the same maximum tolerance (maximum *G-tolerance*).

³ $\chi[T; w_1, \dots, w_n]$ is denoted by χ if there is no confusion about the strict separating system.

⁴From now on we will omit the word *monotonic*.

⁵From now on we will omit the word *linearly*.

Second, we propose a way to get maximum tolerance and maximum G -tolerance.

Theorem 3.3

Maximum tolerance *and* maximum G -tolerance among all equivalent strict separating systems are attained when the weights are natural numbers and their sum is the minimum achievable one.

Given a strict separating system with fixed weights, $[T; w_1, \dots, w_n]$, it is known (Freixas and Molinero, 2008a) that adjusting the corresponding threshold, $\frac{A+B}{2}$ for the tolerance and \sqrt{AB} for the G -tolerance, the system achieves the maximum tolerance and the maximum G -tolerance, respectively. Thus, given a strict separating system of an arbitrary separable switching function, we have a procedure to compute a strict separating system with maximum tolerance or/and maximum G -tolerance, respectively.

For instance, let $[5; 8, 8, 6, 3, 3]$ be a strict separating system, then $\tau = \frac{1}{5+28} = 0.0303\dots$; but the equivalent strict separating natural system with natural weights having minimum sum and threshold equal to $\frac{A+B}{2}$, i.e., $[\frac{3}{2}; 2, 2, 2, 1, 1]$, achieves the maximum available tolerance $\tau' = \frac{1/2}{3/2+8} = 0.0526\dots$. In the same vein, the strict separating system $[5; 8, 8, 6, 3, 3]$ has $\chi = \frac{1}{11} = 0.0909\dots$, but the equivalent strict separating natural system with natural weights having minimum sum and threshold equal to \sqrt{AB} , i.e., $[\sqrt{2}; 2, 2, 2, 1, 1]$, achieves the maximum available G -tolerance $\chi' = \frac{\sqrt{2}-1}{\sqrt{2}+1} = 0.1715\dots$

Another important new result is how we get the maximum tolerance with n variables:

Proposition 3.4

The maximum tolerance for n variables is given by the following strict separating (natural) system:

$$[\frac{1}{2}; 1, \dots, 1],$$

and such tolerance is $\tau = \frac{1}{1+2n}$.

Any other strict (natural or not) separating system (of non null weights) with tolerance τ' fulfills $\tau' \leq \tau$.

Finally, we extend this result for systems with n variables and k distinguished types of variables, i.e., with exactly k non-equivalent variables:

Conjecture 3.5

The maximum achievable tolerance for n variables and k types of distinguished variables is obtained by the following strict separable system:

$$[k - \frac{1}{2}; k, k - 1, k - 2, \dots, 1, \underbrace{1, \dots, 1}_{n-k}]$$

and such a tolerance is $\tau = \frac{1}{2n+k^2+k-1}$.

Again, any other strict linearly (natural or not) separating system (of non null weights) with tolerance τ' fulfills $\tau' \leq \tau$.

In the same vein, we establish the corresponding conjecture for the G -tolerance:

Conjecture 3.6

The maximum G -tolerance for n variables and k types ($k > 1$) of distinguished variables is given by the following strict separable system:

$$[\sqrt{k(k-1)}; k, k-1, k-2, \dots, 1, \underbrace{1, \dots, 1}_{n-k}].$$

and such a G -tolerance is $\chi = \frac{\sqrt{k(k-1)} - (k-1)}{\sqrt{k(k-1)} + (k-1)}$.

In general, given a strict separating system for an arbitrary separable switching function, we are able to find an equivalent strict separating system (with natural weights having minimum integer sum) achieving the maximum tolerance and the maximum G -tolerance. Moreover, it is known, see (Freixas and Molinero, 2008b), that for less than 8 variables all linearly separable switching functions have, up to isomorphism, a unique strict separating system with natural weights having minimum sum. So, we only need to adjust the threshold, $\frac{A+B}{2}$ for the tolerance and \sqrt{AB} for the G -tolerance, for these systems to achieve the maximum tolerance and the maximum G -tolerance, respectively. Note that for 8 variables the situation changes: There are 154 linearly separable switching functions, up to isomorphism, with two strict separating systems having natural weights and minimum sum.

4 Future Work

Generating a huge number of random strict separating natural systems and then to conjecturate what distribution follows the tolerance, the maximum tolerance, the G -tolerance and the maximum G -tolerance.

It is also interesting to provide tables of all, up to isomorphism, linearly separable switching functions of a reasonable high number of variables with: a strict separating system achieving its tolerance; its tolerance; a strict separating system achieving its G -tolerance; and its G -tolerance.

Although the computational limitation (complexity) of finding the tolerance and the G -tolerance of a given strict separating system (both are NP-hard), it will be of interest to develop an efficient algorithm able to calculate the tolerance and the greatest tolerance for strict separating systems with a reasonable high number of variables.

Acknowledgements

J. Freixas was partially supported by Grant MTM 2006–06064 of “Ministerio de Educación y Ciencia y el Fondo Europeo de Desarrollo Regional”, and SGRC 2009-1029 of “Generalitat de Catalunya”.

X. Molinero were partially supported by the Spanish “Ministerio de Educación y Ciencia” programme TIN2006-11345 (ALINEX), and SGRC 2009-1137 of “Generalitat de Catalunya”.

J. Freixas and X. Molinero were also partially supported by Grant MTM2009-08037 of “Ministerio Ciencia e Innovación”, and Grant 9-INCREC-11 of “(PRE / 756 / 2008) Ajuts a la iniciació/reincorporació a la recerca (Universitat Politècnica de Catalunya)”.

References

- Beiu, V., Quintana, J. M., and Avedillo, M. J. (2003). Vlsi implementation of threshold logic—a comprehensive survey. *IEEE Transactions on Neural Networks*, 14(5):1217–1243.
- Dertouzos, M. L. (1965). *Threshold Logic: A Synthesis Approach*. MIT Press.
- Elgot, C. C. (1961). Truth functions realizable by single threshold organs. In *AIEE Conference Paper 60-1311 (October), revised November 1960; paper presented at IEEE Symposium on Switching Circuit Theory and Logical Design*.
- Elizondo, D. (2004). Searching for linearly separable subsets using the class of linear separability method. In *Proceedings of the 2004 International Joint Conference on Neural Networks (IJCNN 04)*, volume 2 of *IEEE*, pages 955–959.
- Elizondo, D. (2006). The linear separability problem: some testing methods. *IEEE Transactions on Neural Networks*, 17(2):330–344.
- Freixas, J. and Molinero, X. (2008a). The greatest allowed relative error in weights and threshold of strict separating systems. *IEEE Transactions on Neural Networks*, 19(5).
- Freixas, J. and Molinero, X. (2008b). On the existence of a minimum integer representation for weighted voting systems. *Annals of Operations Research*, 166(1):243–260.
- Hammer, P. L., Ibaraki, T., and Peled, U. N. (1981). Threshold numbers and threshold completions. *Ann. Discrete Math.*, 11:125–145.
- Hu, S. T. (1960). Linearly separable switching functions. Technical report, Technical report, LMSC, Technical Document, LMSD-703024.
- Hu, S. T. (1965). *Threshold Logic*. Univ. of California Press. xiv + 338 pp. Let xx, x2.
- Muroga, S. (1971). *Threshold logic and its applications*. Wiley-Interscience, New York, USA.
- Myhill, J. and Kautz, W. H. (1961). On the size of weights required for linear–input switching functions. *IRETransactions*, EC-10.
- Ozdemir, H., Kepkep, A., Pamir, B., Leblebici, Y., and Cilingiroglu, U. (1996). A capacitive threshold logic gate. *IEEE J. Solid–State Circuits*, 31.
- Parberry, I. (1994). *Circuit Complexity and Neural Networks*. The M.I.T. Press.
- Picton, P. D. (1991). *Neural Networks*. Palgrave, second edition.
- Roychowdhury, V., Siu, K., and (Eds.), A. O. (1994). *Theoretical Advances in Neural Computation and Learning*. Kluwer Academic Publishers, Stanford, USA.

- Sang-Hoon, O. and Lee, Y. (1995). Sensitivity analysis of single hidden-layer neural networks with threshold functions. *IEEE Trans. Computers*, 6(4):1005–1007.
- Siu, K., Roychowdhury, V., and Kailath, T. (1995). *Discrete Neural Computation: A Theoretical Foundation*. Prentice Hall, New Jersey, USA.
- Winder, R. O. (1962). *Threshold logic*. PhD thesis, Department of Mathematics, Princeton University.
- Yao, S. S. and Ostapko, D. L. (1968). Realization of a class switching functions by threshold-logic networks. *IEEE Transactions on Computers*, C-17(4):391–399.