A Characterization of Hard-threshold Boolean Functions

Fernanda Botelho*

Department of Mathematical Sciences, The University of Memphis, Memphis, TN

This paper characterizes a class of *n*-valued boolean functions, designated as type hard-threshold, in terms of their action on the underlying space graph.

1. Introduction

Boolean functions are discrete-valued functions that provide iterative models in different areas of science. They appear in theoretical computer science as automata networks, cellular, and threshold automata [2–4], and also in biomathematics as input-output functions of perceptrons and Hopfield networks [1, 5, 10]. We can also find them in physics as discrete models for disordered matter such as the spin glass problem [6].

In this paper we consider a class of *n*-valued boolean functions, designated hard-threshold, that represent the activity of discrete recurrent neural network (NN) models.

Recurrent NNs are mathematical models consisting of a large number of parallel-operating and interconnected processing units. Information is encoded as a binary sequence (designated state vector) and is distributed among the units composing the network. This permits a high speed more economical machine implementation and avoids the expensive maintenance of a central database unit [12].

The activity of a network consists of an interchange of information stored in the units composing the net. Based on the information flow and the present state of each unit (input), the network makes a synchronous update of its state vector, the new vector is designated an output.

Hard-threshold boolean functions represent the dynamics of networks that mimic specific brain activity such as associative memory [5]. Examples of such networks are those that associate to the picture of an animal its name or to a picture of a car its model.

Given pairs of binary sequences (x,y), the action of an associative memory network on a given input x should result in the corresponding

^{*}Electronic mail address: botelhof@msci.memphis.edu.

output y. Moreover, given a corrupted or incomplete version of the input, the network should still recover the expected output y.

Minsky and Papert introduced the positive normal form of a boolean function (Theorem 1.5.1 in [7]). This is a decomposition of the function into a linear combination of predicates, canonically defined. The set of predicates involved depends intrinsically on the function considered. We designate as *hard-threshold* a boolean function that is linear threshold in each component, relative to a specific set of predicates [7]. We provide a characterization of those boolean functions that are of type hard-threshold. This characterization is based on the action of the map on the vertices of its underlying graph.

In section 2 we present and motivate the notation and basic definitions used throughout the paper, in section 3 we state and prove the main result.

2. Basic definitions and notation

The basic structure of a Hopfield network with n units is shown in Figure 1. Each unit or cell is connected to each of the other units composing the net with a channel, that represents a real synapsis between two neural cells. Each cell may also be connected to itself via a feedback connection. Signals passing through a network connection change linearly by a multiplicative factor, called weight. The weight matrix collects all of these multiplicative factors and is denoted by \mathbf{W} .

The incoming signal to a given cell is just the sum of all linearly altered signals sent from any cell connected to that specific one. As soon as a signal reaches a cell, it is modified by an external input (i.e., the corresponding component of a constant vector $\Theta = \{\theta_i\}_{i=1..n!}$) and by a transfer function. The transfer function represents a cell response to a signal. A cell may send a signal down its axon or remain passive, depending on the intensity of the incoming signal. This is represented by the function σ that at each real number x assigns 1 if $x \ge 0$ and 0 otherwise. A state vector is a binary n-sequence whose components represent the activation states of the network cells.

We now set notation to be followed throughout the paper. Boldface lower case letters represent state vectors of the network which are vertices of the unit hypercube $[0,1]^n$. A point \mathbf{x} is an n-tuple of 0s and 1s, $\{x_i\}_{i=1}^n$, with x_i representing the activation state of cell i. The connecting weight ω_{ij} is attached to the connection from cell j to cell i and is the ij entry of the connecting matrix \mathbf{W} . The action of \mathbf{W} on some point \mathbf{x} is given by the standard matrix multiplication and denoted by $\mathbf{W}(\mathbf{x})$ or just $\mathbf{W}\mathbf{x}$. For simplicity of notation we define $\sigma(\mathbf{x})$ to be the column vector $[\sigma(x_i)]_i$. The map π_j represents the standard projection onto the jth component.

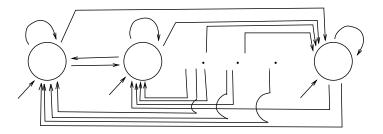


Figure 1. Hopfield network with *n* cells.

The set of all *n*-sequences of 0s and 1s is denoted by \mathcal{V} . A boolean map is a transformation on \mathcal{V} . Definition 2.1 introduces the concept of a boolean map of type hard-threshold which is related to the definition of threshold linearity introduced in [7]. More precisely, T is of type hard-threshold if, for every $i = 1, \ldots, n, \pi_i(T)$ is threshold linear with respect to the set of predicates $\Phi = {\pi_1, \pi_2, \ldots, \pi_n}$, where π_i is the projection on the *i*th component.

Definition 2.1. A boolean map T is of type hard-threshold if there exists a matrix W and a constant n-vector Θ such that

$$T(\mathbf{x}) = \sigma(\mathbf{W}\mathbf{x} - \Theta).$$

Since \mathcal{V} is a finite set, for a given matrix \mathbf{W} , we can always choose Θ so that $\mathbf{W}\mathbf{x} - \Theta \neq 0$ for all $\mathbf{x} \in \mathcal{V}$. This assumption is made throughout the paper.

Definition 2.2. (See [8]) Given x and y points in V, y is called an *immediate neighbor* of x if and only if there exists $i \in \mathbb{N}$ such that

$$\mathbf{x}_k = \mathbf{y}_k$$
 for $k \neq i$ and $\mathbf{x}_i \neq \mathbf{y}_i$.

For simplicity of notation and to emphasize the dependence on the *i*th coordinate, we represent the immediate neighbor of \mathbf{x} altered at site *i* by \mathbf{x}^i . The set of all immediate neighbors of \mathbf{x} is denoted by $\mathcal{N}_{\mathbf{x}}$.

Definition 2.3. A subset S of V is called *connected* if and only if for every x and y in S there exists a sequence in S, $\{x^0 = x, x^1, ..., x^j = y\}$, such that x^k is an immediate neighbor of x^{k-1} .

Associated to a boolean function T, we consider the following two sets at level i.

- 1. The positive set $\mathcal{P}_i = \{\mathbf{x} \in \mathcal{V} : T(\mathbf{x})_i = 1\}.$
- 2. The negative set $N_i = \{ \mathbf{x} \in \mathcal{V} : T(\mathbf{x})_i = 0 \}$.

We also denote the convex hull of \mathcal{P}_i and \mathcal{N}_i , in \mathbb{R}^n , by $\langle \mathcal{P}_i \rangle$ and $\langle \mathcal{N}_i \rangle$, respectively.

3. The main result

Proposition 3.1. A boolean map T is of type hard-threshold if and only if for every i, \mathcal{P}_i and \mathcal{N}_i are connected sets in \mathcal{V} and $\langle \mathcal{P}_i \rangle \cap \langle \mathcal{N}_i \rangle = \emptyset$.

We first prove Lemmas 3.1 and 3.2 that will be used in the proof of Proposition 3.1. These lemmas are stated for a boolean function T of type hard-threshold. Given a point $\mathbf{x} \in \mathcal{V}$, we consider the set $\mathbf{P}_{\mathbf{x}} = \{i : \mathbf{x}_i = 1\}$.

Lemma 3.1. If $\mathbf{x} \in \mathcal{P}_i$, $\mathbf{y} \in \mathcal{N}_i$, and \mathbf{y} is an immediate neighbor of \mathbf{x} , then $\omega_{ik} > 0$ if $k \in \mathbf{P}_{\mathbf{x}}$ and $\omega_{ik} < 0$ if $k \notin \mathbf{P}_{\mathbf{x}}$.

Proof. We assume that $y = x^k$ (using the notation described in Definition 2.2). We have that

$$T(\mathbf{y})_i = \sum_{j=1}^n \omega_{ij} \mathbf{y}_j - \theta_i = \left\{ \begin{array}{ll} \sum_{j \in \mathbf{P_x}} \omega_{ij} - \omega_{ik} - \theta_i < 0 & \text{if} \quad k \in \mathbf{P_x} \\ \sum_{j \in \mathbf{P_x}} \omega_{ij} + \omega_{ik} - \theta_i < 0 & \text{if} \quad k \notin \mathbf{P_x}. \end{array} \right.$$

The statement in the lemma follows from the assumption that $\sum_{j \in \mathbf{P_x}} \omega_{ij} - \theta_i > 0$.

Given a subset S of V we denote by N_S the set of all points that are immediate neighbors of some point in S. The backslash (\) that appears in the statement of Lemma 3.2 represents the difference between two sets.

Lemma 3.2. If S is a connected subset of P_i (or N_i) such that $N_S \setminus S \subset N_i$ (or P_i) then $N_{N_c \setminus S} \setminus S \subset N_i$ (or P_i , respectively).

Proof. Let $\mathbf{z} \in \mathcal{N}_{\mathcal{N}_S \setminus S} \setminus \mathcal{S}$, $\mathbf{y} \in \mathcal{N}_S \setminus \mathcal{S}$ such that $\mathbf{y}^k = \mathbf{z}$, and $\mathbf{x} \in \mathcal{S}$ such that $\mathbf{y} = \mathbf{x}^t$ ($t \neq k$). Therefore, we have that for $i \neq k$ and $i \neq t$, $\mathbf{z}_i = \mathbf{y}_i = \mathbf{x}_i$, $\mathbf{z}_k \neq \mathbf{y}_k = \mathbf{x}_k$, and $\mathbf{z}_t = \mathbf{y}_t \neq \mathbf{x}_t$. This implies that

$$\sum_{i=1}^n \omega_{ij} z_j - \theta_i = \left\{ \begin{array}{ll} \sum_{j=1}^n \omega_{ij} \mathbf{y}_j + \omega_{ik} - \theta_i & \text{if} \quad \mathbf{y}_k = 0 \, (\Leftrightarrow \, k \notin \mathbf{P_x}) \\ \sum_{j=1}^n \omega_{ij} \mathbf{y}_j - \omega_{ik} - \theta_i & \text{if} \quad \mathbf{y}_k = 1 \, (\Leftrightarrow \, k \in \mathbf{P_x}). \end{array} \right.$$

Moreover, $\mathbf{y} \in \mathcal{N}_i$ and Lemma 3.1 imply that $\sum_{j=1}^n \omega_{ij} z_j - \theta_i < 0$. The second case stated in the lemma follows by similar arguments.

Proof of Proposition 3.1. First, we assume T is of type hard-threshold with \mathcal{P}_i and \mathcal{N}_i the positive and negative sets at level i, respectively. If \mathcal{P}_i is not connected then Lemma 3.2 applied to a connected component leads to a contradiction. Similarly if \mathcal{N}_i is not connected. This shows

that both \mathcal{P}_i and \mathcal{N}_i are connected. Furthermore, Definition 2.1 implies that their convex hulls are disjoint.

Conversely, we assume that for every i, \mathcal{P}_i and \mathcal{N}_i are connected with disjoint convex hulls. There are points $\mathbf{x} \in \langle \mathcal{P}_i \rangle$ and $\mathbf{y} \in \langle \mathcal{N}_i \rangle$ whose distance is the shortest possible, that is, $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = d(\langle \mathcal{P}_i \rangle, \langle \mathcal{N}_i \rangle)$, where d represents the usual euclidean distance in R^n and $\|.\|$ the standard norm. Let $\alpha = (\mathbf{x} + \mathbf{y})/2$, the middle point between \mathbf{x} and \mathbf{y} . Now we show that \mathbf{x} is unique, similar arguments prove that \mathbf{y} is also unique. Suppose there exists $\mathbf{x}_1 \in \langle \mathcal{P}_i \rangle$ different from \mathbf{x} such that $\|\mathbf{x}_1 - \alpha\| = \|\mathbf{x} - \alpha\|$, then for $0 < \lambda < 1$ we have

$$\begin{split} \|\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x} - \alpha\|^2 &= \lambda^2\|\mathbf{x}_1 - \alpha\|^2 + (1-\lambda)^2\|\mathbf{x} - \alpha\|^2 \\ &+ 2\lambda(1-\lambda)(\mathbf{x} - \alpha, \, \mathbf{x}_1 - \alpha). \end{split}$$

The expression $(\mathbf{x} - \alpha, \mathbf{x}_1 - \alpha)$ refers to the usual inner product in \mathbb{R}^n . Schwartz inequality (cf. [11]) implies that

$$\|\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x} - \alpha\|^2 \begin{cases} = \|\mathbf{x} - \alpha\|^2 & \text{if} \quad \mathbf{x} - \alpha = \mu(\mathbf{x}_1 - \alpha) \text{ or} \\ & \mathbf{x}_1 - \alpha = \mu(\mathbf{x} - \alpha) \end{cases}$$

$$< \|\mathbf{x} - \alpha\|^2 & \text{otherwise.}$$

Clearly $\|\mathbf{x}_1 - \alpha\|^2 < \|\mathbf{x} - \alpha\|^2$ cannot occur since $d(\langle \mathcal{P}_i \rangle, \langle \mathcal{N}_i \rangle) = 2\|\mathbf{x} - \alpha\|$. If $\mathbf{x} - \alpha = \mu(\mathbf{x}_1 - \alpha)$ or $\mathbf{x}_1 - \alpha = \mu(\mathbf{x} - \alpha)$ then $\mu = 1$ or -1. Both cases are impossible. In fact, if $\mu = 1$ then $\mathbf{x} = \mathbf{x}_1$, if $\mu = -1$ then $\alpha = (\mathbf{x} + \mathbf{x}_1)/2 \in \langle \mathcal{P}_i \rangle$. By a shift of the coordinate system we assume that the origin coincides with α and we represent by $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ the vectors $\alpha \vec{\mathbf{x}}$ and $\alpha \vec{\mathbf{y}}$, respectively.

We define a linear transformation on R^n :

$$f(\vec{\mathbf{z}}) = \frac{(\vec{\mathbf{z}}, \, \vec{\mathbf{x}})}{||\vec{\mathbf{x}}||}.$$

For simplicity of notation we represent a vector $\vec{\alpha}\mathbf{z}$ by \mathbf{z} . We prove that f is positive on $\langle \mathcal{P}_i \rangle$ and negative on $\langle \mathcal{N}_i \rangle$. The point $\mathbf{x} \in \mathcal{P}_i$ and $f(\mathbf{x}) = ||\mathbf{x}|| > 0$. If there exists a point $\mathbf{z} \in \mathcal{P}_i$ such that $f(\mathbf{z}) = 0$ then \mathbf{z} is orthogonal to \mathbf{x} . Given $0 < \lambda < 2||\mathbf{x}||^2/(||\mathbf{z}||^2 + ||\mathbf{x}||^2)$, we have that

$$||\lambda z + (1-\lambda)\mathbf{x}||^2 < ||\mathbf{x}||^2$$

which implies that $d(\langle \mathcal{P}_i \rangle, \langle \mathcal{N}_i \rangle) < \|\mathbf{x} - \mathbf{y}\|$. Similar considerations show that f is negative on \mathcal{N}_i .

We define the hyperplane $\mathbf{H} = \{\mathbf{z} : f(\mathbf{z}) = 0\}$. This hyperplane separates the sets \mathcal{N}_i and \mathcal{P}_i . T is of type hard-threshold with connecting weights given by $\omega_{ij} = (x_i - y_i)/2$ and external input $\theta_i = (x_i^2 - y_i^2)/4$.

Example 1. This example shows that the disjointness required in Proposition 3.1 is in fact necessary. We consider a boolean map T defined on $\{0, 1\}^3$ such that

$$\mathcal{P}_1 = \{(0, 0, 1), (0, 1, 1), (0, 1, 0), (1, 1, 0)\}$$

and

$$\mathcal{N}_1 = \{0, 0, 0\}, (1, 0, 0), (1, 0, 1), (1, 1, 1)\}.$$

We verify that this map is not of type hard-threshold. Suppose otherwise that for every \mathbf{x} , $\pi_1(T(\mathbf{x})) = \sigma(\sum_{i=1}^3 \omega_{1i} x_i - \theta_1)$. Since $\pi_1(T(0, 0, 0)) = 0$, we have that $\theta_1 > 0$. Furthermore, $\pi_1(T(1, 1, 1)) = 0$ and $\pi_1(T(1, 1, 0)) = 1$ imply that $\omega_{13} < 0$ and $\pi_1(T(0, 0, 1)) = 0$ which contradicts that $(0, 0, 1) \in \mathcal{P}_1$. It is easy to check that both \mathcal{P}_1 and \mathcal{N}_1 are connected subsets of $\{0, 1\}^3$ and that $(1/2, 1/2, 1/2) \in \langle \mathcal{P}_1 \rangle \cap \langle \mathcal{N}_1 \rangle$.

We now introduce a definition of a path between two points in V and between two subsets of V.

Definition 3.1. Given two points in \mathcal{V} , \mathbf{x} and \mathbf{y} , we define a *path* between them to be a finite sequence $\{\mathbf{z}^0 = \mathbf{x}, \mathbf{z}^1, \dots, \mathbf{z}^{k-1}, \mathbf{z}^k = \mathbf{y}\}$ such that \mathbf{z}^i is an immediate neighbor of \mathbf{z}^{i-1} . A path between two sets is a path between two points, one point in each set. The *length* of a path is the number of elements in the sequence defining the path.

Proposition 3.2. For every *i* such that $\langle \mathcal{P}_i \rangle \cap \langle \mathcal{N}_i \rangle = \emptyset$, \mathcal{P}_i and \mathcal{N}_i are connected.

Proof. We assume that \mathcal{P}_i is not connected, then it has at least two connected components. We select two components of \mathcal{P}_i , \mathcal{C}_1 and \mathcal{C}_2 , whose path between has the shortest possible length. We represent such a path by the sequence

$$\phi$$
: { \mathbf{x}^1 , \mathbf{x}^2 , \mathbf{x}^3 , ..., \mathbf{x}^k },

where $\mathbf{x}^1 \in C_1$, $\mathbf{x}^k \in C_2$, and $\mathbf{x}^j \in \mathcal{N}_i$ $(j \neq 1 \text{ and } k)$. We associate to this path an injective sequence of indices $\{i_1, i_2, ..., i_{k-1}\}$, where i_t is such that

$$\mathbf{x}_{i_t}^t \neq \mathbf{x}_{i_t}^{t+1}$$
 and $\mathbf{x}_p^t = \mathbf{x}_p^{t+1}$, if $p \neq i_t$.

We consider a new path between x^1 and x^k ,

$$\mathbf{x}^1$$
, $\overline{\mathbf{x}}^2$, $\overline{\mathbf{x}}^3$, ..., $\overline{\mathbf{x}}^{k-1}$, \mathbf{x}^k ,

defined as:

The point
$$\overline{\mathbf{x}}^{t+1}$$
 is such that $\overline{\mathbf{x}}_{i_{k-t}}^{t+1} \neq \overline{\mathbf{x}}_{i_{k-t}}^{t}$ and $\overline{\mathbf{x}}_{j}^{t+1} = \overline{\mathbf{x}}_{j}^{t}$, for $j \neq i_{k-t}$.

The points $\overline{\mathbf{x}}^2$, ..., $\overline{\mathbf{x}}^{k-1}$ are in \mathcal{N}_i since ϕ is of shortest length. Therefore $(\mathbf{x}^1 + \mathbf{x}^k)/2 = (\overline{\mathbf{x}}^2 + \overline{\mathbf{x}}^{k-1})/2 \in \langle \mathcal{P}_i \rangle \cap \langle \mathcal{N}_i \rangle$. This leads to a contradiction which proves the statement.

Propositions 3.1 and 3.2 are summarized in Theorem 3.1.

Theorem 3.1. A boolean map is of type hard-threshold if and only if for every i, $\langle \mathcal{P}_i \rangle \cap \langle \mathcal{N}_i \rangle = \emptyset$.

Remark

If we assign 1 to an even integer and 0 otherwise, given a sequence of n integers we associate its corresponding binary sequence. The value of T on a binary n-sequence, $\mathbf{x} = \{x_i\}_{i=1,\dots,n}$ is given by $T(\mathbf{x}) = \sigma\left(\sum_{i=1}^n x_i - 1/2\right)$. The map T defines a partition of the vertices of the hypercube $[0,1]^n$ into two disjoint sets. The set $N = \{0\}_{i=1,\dots,n}$ and its complement P. These two sets clearly satisfy the conditions of Theorem 3.1. Therefore a NN with dynamics given by T is capable of deciding if a product of n integers is even or odd.

Acknowledgements

This work was done while the author was visiting the Institute for Mathematics and Its Applications (IMA) at the University of Minnesota. IMA is supported by the National Sciences Foundation. The author gratefully acknowledges the constructive remarks of the reviewer.

References

- [1] Amit, D., Modeling Brain Function, the World of Attractor Neural Networks (Cambridge University Press, Cambridge, 1989).
- [2] Albert, J. and Culik, K., "A Simple Universal Cellular Automata and Its One-way and Totalistic Version," *Complex Systems*, 1 (1987) 1–16.
- [3] Codd, E., Cellular Automata (Academic Press, New York, 1968).
- [4] Garzon, M., Models of Massive Parallelism: Analysis of Cellular Automata and Neural Networks (Springer-Verlag, Germany, 1995).
- [5] Hopfield, J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proceedings of the National Academy of Sciences of the USA*, 79 (1982) 2554–2558.
- [6] Mezard, M., Parisi, G., and Virasoro, M., *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
- [7] Minsky, M. and Papert, S., *Perceptrons: An Introduction to Computational Geometry* (MIT Press, Cambridge, MA, 1969).
- [8] Robert, F., Discrete Iterations (Springer-Verlag, Berlin, 1986).
- [9] Rosenblatt, F., "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Psychological Review*, **65** (1958) 386–408.

[10] Rosenblatt, F., *Principles of Perceptrons* (Spartan, Washington D.C., 1962).

- [11] Rudin, W., *Principles of Mathematical Analysis* (McGraw-Hill, New York, 1976).
- [12] De Wilde, P., Neural Network Models (Springer-Verlag, Great Britain, 1997).