Journal of Pure and Applied Algebra 83 (1992) 197–203 North-Holland 197

Binary classifiers, perceptrons and connectedness in metric spaces and graphs

Benedict G. Seifert

Corpus Christi College, Oxford University, St. Giles, Oxford OX1 4JW, United Kingdom

Communicated by J. Rhodes Received 15 May 1990 Revised 10 February 1992

Abstract

Seifert, B.G., Binary classifiers, perceptrons and connectedness in metric spaces and graphs, Journal of Pure and Applied Algebra 83 (1992) 197–203.

Minsky and Papert's well-known results about the inability of diameter-limited perceptrons to recognise connectedness in the plane are generalised in three important respects: First, to a large class of metric spaces of arbitrary dimension, not necessarily Euclidean; secondly, to the class of all connected graphs; thirdly, in the case of infinite diameter, including Minsky and Papert's case, it is shown that no diameter-limited binary classifier (including neural nets of arbitrary complexity, etc.) can recognise connectedness.

1. Introduction

In Minsky and Papert's important book 'Perceptrons', the authors give a large variety of results about the computational features of certain classes of 1st order perceptrons. These classes are defined by the 'sensors' which form the 0th order computational device which, when composed with the linear-threshold function of level 1, defines the perceptron. Typically, the 0th level device is an arbitrary function on subsets of given diameter of some metric space. Some of these results show, sometimes constructively, that these classes of perceptrons can compute the characteristic function of classes of sets defined by certain geometric properties. Other results show that these classes do not contain the characteristic function of such classes. The most famous example is the result that diameter-limited

Correspondence to: Dr. B.G. Seifert, Corpus Christi College, Oxford University, 26 Worcester Place, St. Giles, Oxford OX1 4JW, United Kingdom.

0022-4049/92/\$05.00 © 1992 - Elsevier Science Publishers B.V. All rights reserved

perceptrons cannot compute the predicate 'S is connected' for subsets of the plane, both ordinary and discretised.

In this paper we generalise Minsky and Papert's results in two respects. First, we generalise the types of metric spaces to which such results apply, both for the continuous plane (Section 3) and the discretised plane (Section 4). The generalisation from the ordinary plane leads us to define connected *r*-connected metric spaces, which are connected in the standard sense of point-set topology and all of whose discs of radius $\leq r$ are also connected. We generalise Minsky and Papert's results concerning the discretised plane to all symmetric connected graphs. In this context the notion of 'connected set' means pathwise connected, where paths are simply sequences of vertices such that successive pairs of adjacent points along the path are edges.

Furthermore, in the case of spaces of infinite diameter, we give a very strong generalisation of Minsky and Papert's result on connectedness: We generalise their theorem to all binary classifiers operating on diameter-limited sensors. It thus turns out that this result has nothing whatever to do with perceptrons and indeed with neural networks of arbitrary architecture, but uses only the most general and obvious properties of the class of connected subsets of graphs and continuous metric spaces of the sort envisaged in this paper.

We now give a summary of individual sections. In Section 2, we introduce a simple result (Fundamental Lemma 2.2), important in the sequel, on classes of subsets in metric spaces recognised by perceptrons. In Section 3, we prove the generalisation of Minsky and Papert's results to 1st order perceptrons defined on certain continuous connected metric spaces. In Section 4, we prove the analogous results for connected symmetric graphs. In Section 5, we give the generalisation to arbitrary diameter-limited binary classifiers. In Section 6, we explain the precise relationship between our results and those of Minsky and Papert.

2. The Fundamental Lemma

We shall consider a metric space (X, d) of diameter D. Let M be a class of subsets of X and C a subclass of M. We are interested in the existence of simple methods for 'recognising' C in M computationally, i.e. in the existence of simple binary computational devices which are equal to 1 on S in M iff S is in C and 0 otherwise.

Definition 2.1. (a) Let r > 0. An *r-limited perceptron* is a function *p* from *M* to $\{0, 1\}$ constructed as follows:

(i) $\{S_1, \ldots, S_N\}$ is a family of sets in *M*, each of diameter $\leq r$.

.

(ii) t is a function from $\Re(+)$ to $\{0, 1\}$ defined by:

$$t(x) = \begin{cases} 1 & \text{if } x > 1\\ 0 & \text{if } 0 \le x \le 1 \end{cases}$$

(iii) $\{l_1, \ldots, l_m\}$ are arbitrary functions from M to $\Re(+)$. Given these data, we define the function p from M to $\{0, 1\}$ by

$$p(S) = t\left(\sum_{i=1}^{m} l_i(S \cap S_i)\right).$$
⁽¹⁾

(b) A subclass C of M is said to be *recognised* by the perceptron p iff C consists precisely of the sets S in M for which p(S) = 1.

Fundamental Lemma 2.2. Let (X, d) be a metric space. If M is any class of sets in X, r > 0, p an r-limited perceptron on M, M(p) the class of sets recognised by p, then M(p) has the following additivity property:

If $S, T \in M(p)$ and d(S, T) > r, then $S \cup T \in M(p)$.

Proof. Let $\{S_i\}, \{l_i\}, t$ define p as in Definition 2.1. Then, by hypothesis

$$t\left(\sum_{i=1}^{m} l_i(S \cap S_i)\right) = 1 , \qquad t\left(\sum_{i=1}^{m} l_i(T \cap S_i)\right) = 1 , \qquad (2)$$

and hence

$$\sum_{i=1}^{m} l_i(S \cap S_i) > 1, \qquad \sum_{i=1}^{m} l_i(T \cap S_i) > 1.$$
(3)

We can now define the subsets of the index set $I = \{1, ..., n\}$ as follows:

$$I(T) = \{i \mid l_i(T \cap S_i) \neq 0\},\$$

$$I(S) = \{i \mid l_i(S \cap S_i) \neq 0\},\$$

$$I(0) = \{i \mid l_i(S \cap S_i) = 0 \text{ and } l_i(T \cap S_i) = 0\}.$$
(4)

Since diameter $(S_i) \le r$ and d(S, T) > r, for each *i* at least one of the intersections $S_i \cap S$ or $S_i \cap T$ must be empty. Hence I(T), I(S), I(0) form a partition of *I*. Hence

$$\sum_{i=1}^{m} l_i((S \cup T) \cap S_i) = \sum_{i=1}^{m} l_i(S \cap S_i) + \sum_{i=1}^{m} l_i(T \cap S_i) .$$
(5)

Since, by hypotheses each of the summands in the sum of the right-hand side exceeds the threshold 1, so does their sum. Hence

$$p(S \cup T) = t\left(\sum_{i=1}^{m} l_i((S \cup T) \cap S_i)\right) = 1.$$
(6)

This proves the lemma. \Box

3. Connected *r*-locally connected metric spaces

We consider a class of metric spaces which includes all real and complex Euclidean spaces, all connected Lie groups, spheres and all Riemannian manifolds which are homogeneous spaces under the action of a connected Lie group, and many other examples.

Definition 3.1. A *connected*, *r*-connected metric space is a connected metric space (X, d) all of whose discs of radius $\leq r$ are connected.

We consider M to be the class of all subsets of X, and C the class of all connected subsets. We shall apply the fundamental lemma of the previous section in order to show that the class of connected sets cannot be recognised by r-limited perceptrons unless r is greater than three times the diameter of the metric space.

Theorem 3.2. If the connected, r-connected metric space (X, d) has diameter D, and 2r < D, then C cannot be recognised by an r-limited perceptron.

Proof. Let $D = 2r + \alpha$, $\alpha > 0$. By definition, $D = \sup\{d(x, y) \mid x, y \text{ in } X\}$. For any $\varepsilon > 0$, we can find points x and y such that $d(x, y) > D - \varepsilon > 2r + \alpha - \varepsilon$, and hence the discs $D_r(x)$ and $D_r(y)$, by the triangle inequality, have distance greater than $\alpha - \varepsilon$. Taking $\varepsilon < \alpha$, $d(D_r(x), D_r(y)) > 0$. By definition of r-connectedness, $D_r(x)$ and $D_r(y)$ are connected, i.e. in C. Hence, by the Fundamental Lemma, $D_r(x) \cup D_r(y)$ must be in C, i.e. be connected. But, being the disjoint union of two connected subsets, $D_r(x) \cup D_r(y)$ is not connected. This proves the theorem. \Box

The theorem immediately implies the following:

Corollary 3.3. If (X, d) is of infinite diameter, then the class of connected subsets cannot be recognised by any diameter-limited perceptron. \Box

4. Connected graphs

Let G = (V, E) be a connected symmetric graph, with V the vertices and E the edges of the graph G. By definition, E is a subset of $V \times V$, invariant under the involution i(v, w) = (w, v).

The invariance under *i* insures that the function d(v, w), the length of a shortest path from v to w, defines a metric space structure on V. We let M denote the class of all subsets of V, and C the class of subsets which induce a connected subgraph. Equivalently, C is the class of subsets S such that, for any pair of vertices in S, there exists a path consisting of edges both vertices of which are elements of S.

We record the following immediate result:

Lemma 4.1. All discs $D_r(x)$ are connected. \Box

We then have the following analogue of Theorem 3.2:

Theorem 4.2. Let (V, E) be a symmetric graph, and (V, d) the induced metric space. If r > 0 such that the diameter of V is greater than 2r, then the class C of connected subsets of V cannot be recognised by an r-limited perceptron. \Box

The proof of this theorem is entirely analogous to the corresponding proof of Theorem 3.2.

5. General binary classifiers

Perceptrons are extremely simple binary classifiers. In the preceding sections, we have seen that these classifiers have severe limitations as geometric patternrecognition devices, assuming that individual sensors are diameter-limited. In the present section, we shall show that the restriction to diameter-limited sensors implies similar constraints on arbitrary binary classifiers. In particular, we can generalise the famous result of Minsky and Papert that no diameter-limited perceptron can recognise the property of 'connectedness' in a certain discretised version of the Euclidean plane: We shall show that no binary classifier with diameter-limited sensors, and in particular no such neural network can recognise connected sets in certain connected metric spaces (of which the discretised or continuous plane are simple examples) of infinite diameter.

First we shall define the notion of binary classifiers for metric spaces:

Definition 5.1. Let (X, d) be a metric space, M a family of subsets of X. A *binary diameter-bounded classifier* p on M is a function from M to $\{0, 1\}$ which is the composition $\mu \circ \phi$, where μ and ϕ are defined as follows:

(i) Let S_1, \ldots, S_n be a set of bounded subsets of X. Then ϕ is the function from M to the set B^n of n-tuples of bounded sets in X which assigns to S in M the tuple

$$(S \cap S_1, \dots, S \cap S_n) . \tag{7}$$

(ii) μ is an arbitrary function from B^n to $\{0, 1\}$.

Clearly diameter-limited perceptrons are of this form. An important more general class of binary classifiers can be obtained by letting general feed-forward neural networks of any architecture play the role of μ in this definition.

Definition 5.2. A subfamily C of a family M of sets in (X, d) is said to be of *finite character* iff there exists a subset W of X of finite diameter such that for all subsets S in M, S is in C if and only if $S \cap W$ is.

Lemma 5.3. Let (X, d) be a metric space of infinite diameter, M a class of subsets of X. Let p be a diameter-limited binary classifier on a class M of subsets of X. Then the class of sets in M recognised by p is of finite character.

Proof. It suffices to take W to be the union of the sets S_i in the definition. Then clearly, for any S in M, $\phi(S) = \phi(S \cup W)$, and hence $p(S) = \mu(\phi(S)) = \mu(\phi(S \cap W)) = p(S \cap W)$. \Box

Theorem 5.4. Let M be the class of all subsets of a metric space (X, d) of infinite diameter and satisfying the hypotheses of Section 3 or Section 4, C the class of connected sets in the sense of Section 3 or Section 4. Then C cannot be recognised by any diameter-limited binary classifier.

Proof. It is easy to verify that, in the case of connected graphs (Section 4), as well as that of *r*-connected, connected metric spaces (Section 3) the class of connected subsets is not of finite character. Hence the statement follows from Lemma 5.3. \Box

6. Results of Minsky and Papert

We shall now explain that the results of [2] concerning connectedness in the plane are special cases or direct corollaries of our results.

Minsky and Papert study the pattern recognition capabilities of perceptrons, as defined in this paper, First, they define the notion of 'order' [2, p. 30]. For the purposes of the present exposition, we slightly change this definition and call the order of a predicate defined on a family of subsets of a connected metric space to be the smallest number r such that there exists an r-limited perceptron which can recognise the predicate in question. The reader will be left to verify that our conclusions apply also to the definition adopted by [2].

The authors then give a construction of a 'discrete plane' *P*. *P* consists of the set of squares $\{[i, i+1] \times [j, j+1] | i, j \text{ integers}\}$ in the ordinary Euclidean plane. A graph G = (P, E) is introduced by declaring a pair of squares to be in *E* if two of their boundaries coincide. The resulting graph is connected. Connected sets in the sense of [2] are then just connected sets in the sense of graph theory. Hence our results apply. The authors study *G* as the inductive limit of connected subgraphs $G_i = \{(P_i, E_i) | i = 1, 2, ...\}$ (for example the graphs induced by the subsets $[-i, i] \times [-i, i]$). Given such an inductive family, they study the orders 1_i of the predicate: *S* is a connected set in G_i . They then prove [2, Theorem 5.1, p. 74] the following: **Theorem 6.1.** $\{l_i \mid i = 1, 2, ...\}$ converge to infinity.

Proof. Since clearly the sequence 1_i is monotone increasing, the statement is equivalent to the statement that the numbers 1_i do not converge to a finite limit *l*. Supposing the limit *l* existed, then *l* would also be the order for the predicate 'S in G is connected'. Theorem 3.2 of the present paper implies that *l* cannot exist since the predicate 'S in G is connected' cannot be recognised by a diameter-limited perceptron. \Box

Furthermore, we can strengthen Theorem 6.1, in the light of Section 5. First, we introduce the notion of 'unrestricted order' of a metric space (X, d) as the smallest diameter l such that there exists an l-limited binary classifier which recognises the predicate connectness.

Theorem 6.2. The sequence of unrestricted orders of the graphs G_i , $\{l_i\}_{i=1,...,n}$, converges to infinity.

Proof. Analogously to Theorem 6.1, the statement follows from Theorem 5.4. \Box

References

[2] M. Minsky and S. Papert, Perceptrons (MIT Press, Cambridge, MA, expanded ed. 1988).

^[1] J. Dugundji, Topology (Allyn & Bacon, Boston, MA, 1966).