

AN ATTEMPT IN MODELLING AUTISM USING SELF-ORGANIZING MAPS

Andrew P. Papliński*

Computer Science and Software Engineering
Monash University, Australia

Lennart Gustafsson

Luleå University of Technology,
Sweden

ABSTRACT

Autism is a developmental disorder in which attention shift impairment and strong familiarity preference are considered to be prime deficiencies. We model these two characteristics of autistic behaviour using Self-Organizing Maps (SOFM).

1. INTRODUCTION

Autism is a developmental disorder first described by Kanner [1], and Asperger [2]. Presently diagnostic criteria according to DSM-IV [3] are grouped into three main categories, namely, impairments in social interaction, impairments in verbal and nonverbal communication, and restricted repertoire of activities and interests. The diagnostic criteria are behavior-based, but a number of biological abnormalities have been connected with autism. For an introduction, see Gillberg and Coleman [4].

There is a general agreement that attentional impairment is commonly seen in autism. This impairment includes joint attention and attention shifts. However there are different opinions whether attentional impairment is a primary cause for other autistic characteristics or is itself secondary to some other autistic characteristic.

Theories on causes of autism, based on properties of artificial neural networks, have been presented by Cohen [5] and Gustafsson [6]. The purpose of this paper is to examine how the attention shift impairment and familiarity preference influence the self-organization of an artificial neural network and to discuss the characteristics of the resulting maps. It will be shown that some, but not all, of these maps exhibit characteristics which may be argued to represent autistic behaviour. A comparison will be made with maps organized when novelty seeking is present.

2. NEURAL NETWORKS USED IN MODELLING

We use Kohonen self-organizing maps [7, 8] to model the attention shift impairment and familiarity preference. For convenience we repeat main points important from the point of view of our presentation.

The authors would like to thank the Swedish Foundation STINT and the Monash University SMURF-2 grant schemes for financial support.

Self-Organizing Maps (SOMs) are competitive neural networks in which neurons are organized in an l -dimensional lattice (grid) representing the **feature space**. Such neural networks perform mapping of a p -dimensional **input space** into the l -dimensional feature space. With respect to the visualisation aspect the dimensionality of the feature space is often restricted to $l = 1, 2$ or 3 .

In Figure 1 we present an example of a self-organizing map consisting of $m = 12$ neurons in which the **input space** is 3-dimensional ($p = 3$) and the **feature space** is 2-dimensional ($l = 2$). The first section of the network is a

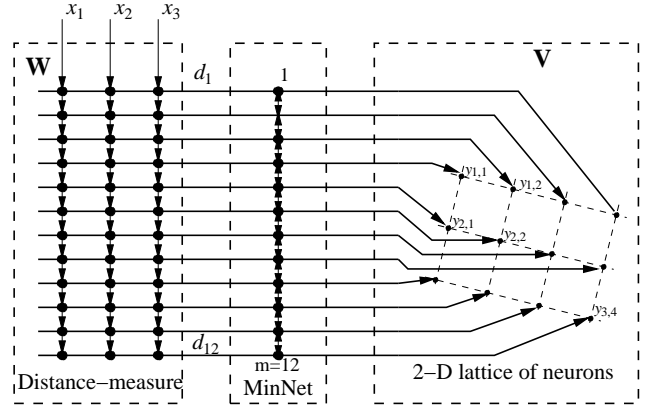


Figure 1: A 2-D SOFM with $p = 3$; $m = [3 \ 4]$; $l = 2$.

distance-measure layer consisting of $m = 12$ dendrites each containing $p = 3$ synapses excited by 3-D input signal vectors $\mathbf{x} = [x_1 \ x_2 \ x_3]$ and characterised by the weight vector $\mathbf{w}_i = [w_{i1} \ w_{i2} \ w_{i3}]$. The distance-measure layer calculates the distances d_i between each input vector \mathbf{x} and every weight vector \mathbf{w}_i . This distance information, (d_1, \dots, d_m) is passed to the competition layer, the MinNet in Figure 1, which calculates the minimal distance $d_k = \min d_i$ in order to establish the position of the winning neuron k . The competition is implemented through the lateral inhibitive and local self-excitatory connections between neurons in the competitive layer. In addition, every neuron is located at $l = 2$ -dimensional lattice and its position is specified by an l -dimensional vector $\mathbf{v}_i = [v_{i1} \ v_{i2}]$.

The synaptic weight vectors, \mathbf{w}_i , and the vectors of topological positions of neurons, \mathbf{v}_i , are grouped into the weight

and position matrices, W , V , respectively.

2.1. Feature Maps

A typical **Feature Map** is a plot of synaptic **weights** in the **input space** in which weights of the neighbouring neurons are joined by lines and illustrates the mapping from p -dimensional input space to l -dimensional feature space. An example of such a map with $p = 2$ inputs and $m = 12$ neurons organized on a 3×4 lattice, together with respective weight W and position V matrices is given in Figure 2.

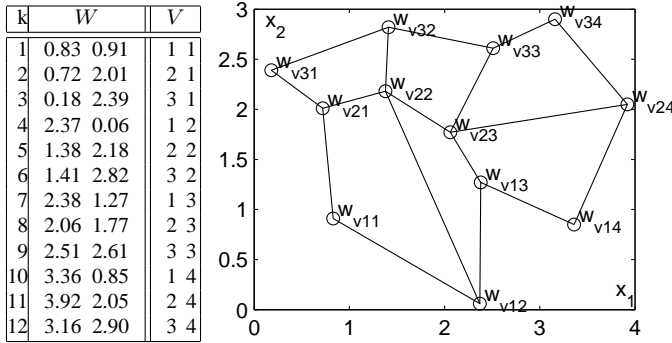


Figure 2: Example of the weight and position matrices and the the resulting feature map for $p, l = 2$

In our modelling we will be using similar SOMs with $p, l = 2$ and neurons organised in either a 2×2 , or 3×3 mesh.

2.2. Map Formation Algorithm

The learning algorithm consists of two essential aspects of the map formation, namely, **competition** and **cooperation** between neurons of the output lattice. During competition each input vector $\mathbf{x}(n)$ is compared with each weight vector from the weight matrix W and the position $V(k(n), :)$ of the winning neuron $k(n)$ is established. For the winning neuron the distance

$$d_k = |\mathbf{x}^T(n) - W(k(n), :)|$$

attains a minimum (notation as in MATLAB [9] used as a computational tool; ‘:’ denotes a complete row of a matrix). During the cooperation phase all neurons located in a topological neighbourhood of the winning neurons $k(n)$ will have their weights updated usually with a strength $\Lambda(j)$ related to their distance $\rho(j)$ from the winning neuron,

$$\rho(j) = |V(j, :) - V(k(n), :)| \quad \text{for } j = 1, \dots, m.$$

The weight matrix W is then modified by

$$\Delta W = \eta(n) \cdot \Lambda(n) \cdot (\mathbf{x}^T(n) - W(j, :))$$

where the **neighbourhood function**, $\Lambda(j)$, is an l -dimensional Gaussian function:

$$\Lambda(j) = \exp\left(-\frac{\rho^2(j)}{2\sigma^2}\right)$$

σ^2 is the variance parameter specifying the spread of the Gaussian function, and η is the learning gain parameter. During the ordering phase, the variance is reduced until the neighbourhood includes only one neuron. During the convergence phase the learning process is being, “cool down” by reducing the learning gain.

3. MODELLING AUTISTIC BEHAVIOUR

In order to model autistic behaviour we arrange the two-dimensional training data, or stimuli, into **two sources**, A and B . Each source contains **three classes** of stimuli specified by a centroid of the class. There are ten exemplars in each class for each source randomly distributed around the centroids.

The sources can be thought of as producing two dialects of a very limited protolanguage, each with three protophonemes. Real sensory stimuli, like phonemes of speech, are of course larger in number and dimension. However, the explanations given for the results obtained do not rely on the dimensionality or the number of classes of the sources. Simulations with sources that provide few classes of low dimensionality have been presented because a complete visualization of the results can easily be achieved.

As a result of training a map similar to that of Figure 2 will be formed. In a “good”, or well developed map, the map nodes representing the weight vectors should coincide with the mean values of each class irrespective of the source. For our data we have three such values, therefore, a canonical (smallest possible) map can have only 2×2 neurons (nodes).

Initial values of weight vectors are located randomly around the total mean of the training data, and during learning the weight vectors will be pulled up towards the class means.

The **map “goodness” index** will be formed from the sum of distances between the class means and weights located in the proximity of the means. This sum of distances will be the largest for the initial value of weights and will be reduced during training. The “goodness index” has been normalized so that a value of 1 represents a map with nodes located at the centres of all data subclasses.

Another important parameter is the **number of attention shifts** measured as a fraction of total stimuli used during learning.

In the **novelty seeking mode of learning (mode 1)** attention is shifted to the alternate source if the next stimulus originates from that source. This is regarded as normal, or

non-autistic mode of learning. An example of maps generated in Mode 1 is given in Figure 3. In this mode the atten-

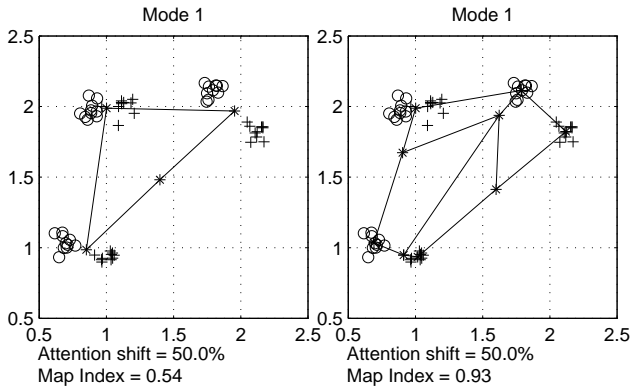


Figure 3: An example of a canonical (2×2) and a 3×3 map in a novelty seeking learning mode.

tion shift fraction is equal to 0.5 which indicates that stimuli originate from alternate sources. The map goodness index for a 3×3 map approaches unity which indicates nearly perfectly spread maps in which the node weights assume values which are the mean values of each class. For a 2×2 map the index is as good as it can get, with nodes adapting to means of the subclasses from both sources combined.

In **mode 2** we test the **attention shift impairment**. In this case attention is shifted to the alternate source with a low probability, in the simulations chosen as 0.01, if that source presents the next new stimulus. The maps resulting from simulation in this mode are presented in Figure 4. It

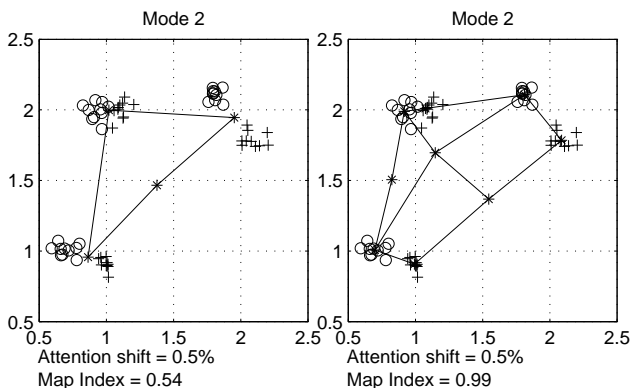


Figure 4: An example of a canonical (2×2) and a 3×3 map in an attention shift impairment learning mode.

might at first seem surprising that the resulting maps from learning with attention shift impairment are the same as those resulting from learning with novelty seeking but the explanation is straightforward — if the nodes were more adapted to the mean values of the subclasses of one of the

sources then learning from exemplars of the other source would result in greater node weight adjustments, pulling the node towards the mean value of the subclasses of both sources. It should be noticed that the number of attention shifts in learning with attention shift impairment is very low (reduced by the probability of attention shift). This factor does not prevent the successful self-organization of the neural network. This is because there is no bias in favor of one of the sources in this mode of learning. Therefore, the map goodness index is similar as in mode 1 (Figure 3).

In **mode 3**, we model **familiarity preference**. In this mode attention is shifted to the alternate source if that source presents the next new stimulus when both sources are unfamiliar to the map, i.e., in the first phase of the self-organization, and then after familiarity to at least one of the sources has been reached with lower probability to the source which is the least familiar to the map. Familiarity of a source to the map is a weighted average of the distances between the weight vectors of the nodes most resembling the stimuli from that source in the past. Maps resulting from simulation in this mode are presented in Figure 5. Note that the

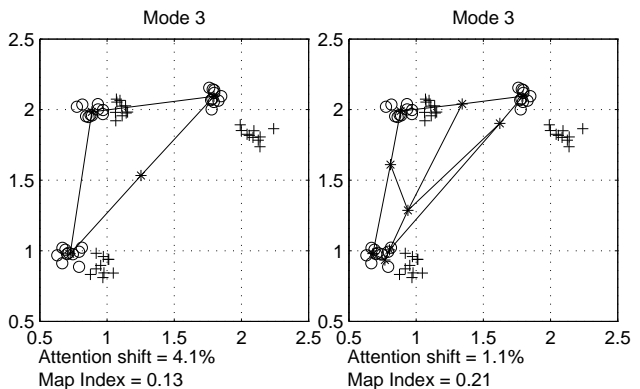


Figure 5: An example of a canonical (2×2) and a 3×3 map in a familiarity preference learning mode.

maps developed in the familiarity preference learning mode tend to give preference to one source. Number of attention shifts however small is higher than in the attention shift impairment mode (Figure 4). The map goodness indexes are smaller indicating the preference of the maps to only one source.

In the next set of simulations one data source is reduced to provide only two classes of stimuli, one of the classes having twenty exemplars. The resulting maps for mode 2 (attention shift impairment) and mode 3 (familiarity preference) are presented in Figure 6. As before, the resulting maps from learning in mode 1 and 2 are somewhat similar and adequately cover all classes of the stimuli as in Figures 3 and 4. The resulting map from learning in mode 3

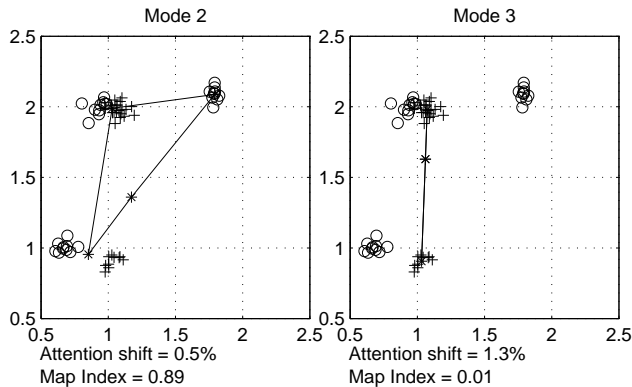


Figure 6: Self Organized maps for a reduced number of classes in one data source.

shows that the source with a reduced set of stimuli dominates the development of learning, leaving one subclass of stimuli from the full source without any detector node.

Again there is a straightforward explanation for this result — the exemplars from the reduced set source show less diversity than the exemplars from the full set source and learning from the exemplars of the reduced set source will therefore cause a faster adaptation of the node weights and thus exemplars from the full set source will subsequently be ignored.

Further details of simulation in mode 3 can be found in [10].

4. CONCLUSION

Our results show that familiarity preference results in inadequate maps with characteristic deficits such that they lend support to the theory that familiarity preference or novelty avoidance may be primary in causing other autistic characteristics. The stimuli of one source will be learned precisely, at the expense of the other. If one source has a reduced set of stimuli it will dominate the resulting map.

If the SOM is canonical this domination will preclude the learning of the source with a full set of stimuli. It may be argued that development of cortical maps of this kind in a child will be inductive of the development of narrow interests, commonly present in autism.

If the SOM has an excess of nodes the result will vary greatly between simulations, even though all initial values are the same for all simulations — in some cases only stimuli from the reduced set source are learned and in others the stimuli from the full set source will also be learned well.

There are many cortical maps and if some of them develop to respond only to a reduced set source and others achieve a normal development this offers a reasonable explanation to the uneven capacities often found in individuals with autism (Kanner’s “islets of ability”; for a discussion,

see e.g. Frith [11]).

Our results do not lend support to the hypothesis that attention shift impairments by themselves may be primary in causing other autistic characteristics since self-organization with attention shift impairments have, in our simulations, always resulted in normal maps.

Learning with attention shift impairments in conjunction with familiarity preference, however, will very much reduce the probability for a normal map resulting from self-organization. Thus attention shift impairment may be important, albeit not independently so, in causing autism.

5. REFERENCES

- [1] L. Kanner, “Autistic disturbances of affective contact.,” *Nervous Child*, no. 2, pp. 217–250, 1943.
- [2] H. Asperger, “Die ‘autistischen Psychopathen’ im Kindesalter.,” *Arch. Psychiatrie Nervenkrankheiten*, no. 117, pp. 76–136, 1944. Translated in Frith U. (ed) (1991): *Autism and Asperger Syndrome*. Cambridge University Press.
- [3] “Diagnostic and statistical manual of mental disorders. 4th ed.,” American Psychiatric Association, 1994. Available from: <http://www.psychology.net/org/dsm.html>.
- [4] C. Gillberg and M. Coleman, *The Biology of the Autistic Syndromes*. Cambridge University Press, 3rd ed., 2000.
- [5] I. Cohen, “An artificial neural network analogue of learning in autism,” *Biol. Psychiatry*, no. 36, pp. 5–20, 1994.
- [6] L. Gustafsson, “Inadequate cortical feature maps: A neural circuit theory of autism,” *Biol. Psychiatry*, no. 42, pp. 1138–1147, 1997.
- [7] T. Kohonen, *Self-Organisation and Associative Memory*. Berlin: Springer-Verlag, 3rd ed., 1988.
- [8] T. Kohonen, *Self-Organising Maps*. Berlin: Springer-Verlag, 2nd ed., 1997.
- [9] MathWorks, *Using MATLAB*. The MathWorks Inc., 2002.
- [10] L. Gustafsson and A. P. Papliński, “An experiment in modelling learning in autism using self-organizing artificial neural networks,” Tech. Rep. 2001/93, School of Comp. Sci. and Soft. Eng., Monash University, Australia, June 2001.
- [11] U. Frith, *Autism: Explaining the Enigma*. Oxford: Basil Blackwell, 1989.