A SOM-Based Validation Approach to a Neural Circuit Theory of Autism

Spyridon Revithis[1] and Georgios Tagalakis[2]

[1] School of Computer Science and Engineering, University of New South Wales, UNSW-Sydney, NSW 2052, Australia
revithiss@cse.unsw.edu.au
[2] School of Computer Science and Informatics, University College Dublin, Dublin 4, Ireland

Abstract. The neural network class of self-organizing maps (SOMs) is a promising cognitive modeling tool in the study of the autistic spectrum pervasive developmental disorder. This work offers a novel validation of Gustafsson's neural circuit theory, according to which autism relates to formation characteristics of cortical brain maps. A previously constructed spatial SOM behavioral model is used here as a cognitive model, and by incorporating formation deficiencies related to the topological neighborhood (TN) function. The resulting cognitive SOM maps, being sensitive to the width of TN during SOM formation, point to a model that exhibits marked behavioral characteristics of autism. The simulation results support the causal hypothesis that associates autistic behavior with certain functional and structural characteristics of the human nervous system and, specifically, Gustafsson's theoretical proposition of the role of inhibitory lateral feedback synaptic connection strengths in autism.

Keywords. Neural Networks, Self-Organizing Maps, Cognitive Modeling, Autism.

1 Introduction

Computational modeling offers a powerful way of studying human behavior. It has been applied to numerous areas of Psychology and provides a framework superior to those proposed by the social sciences in terms of methodological diversity, empirical accuracy, and procedural clarity [1]. An increasing number of studies are dedicated to the modeling of developmental cognitive phenomena using neural networks [2-3].

This study investigates neurocomputational aspects of autism. Section 2 presents some key characteristics of the autistic spectrum disorder and a current neural circuit theory. In Section 3, the details of an autistic SOM model, and the computational simulations performed in order to investigate and evaluate its efficacy, are introduced. Section 4 informs of the model and simulation technical details with supporting statistical evaluation of the results. The concluding Section 5 discusses some planned and other possible directions of future work.

I. Maglogiannis, V. Plagianakos, and I. Vlahavas (Eds.): SETN 2012, LNAI 7297, pp. 25–32, 2012.

2 The Neuropsychology of a Neural Circuit Theory of Autism

Since Kanner's [4] and Asperger's [5] publications, autism, a pervasive developmental disorder, is studied by an ever-expanding interdisciplinary research community. Its etiology remains unknown, but it is considered to be neurobiological in nature [6]. Unfortunately, the current diagnostic tools (DSM-IV and ICD-10) dictate a socio-psychological behavioral approach that does not inform of the causes of autism.

Autism is associated with atypical perception and its internal representation. Sensory input often fails to integrate into existing memory (schemas) due to abstraction impairment [7]. There is difficulty to detect the important features among the non-essential details [8]. Elaborating on internal representations is also problematic, where it appears that central executive control is required [9].

Gustafsson's neural circuit theory of autism [10] is based on these empirically based concepts of autistic perception and proposes a neural-level explanation for the lack of drive for central coherence, a key element in autistic behavior [11]. Specifically, important attributes in autism are derived from neurological deficiencies in the formation of brain cortical maps; this leads to problematic feature extraction since "autistic raw data memory" operates in place of "feature memory" due to "inadequate cortical feature maps". Raw data memory is intrinsically linked at the behavioral level to the diagnostic criteria for autism [10]. Autistic maps lack feature distinction and preservation, and fail to provide an internal representation of salient perceptual data. This leads to raw data memory that lacks sophisticated representations [12].

According to Gustafsson [10], the artificial neural network class of self-organizing maps (SOMs) [13-14] provides a biologically plausible way to model characteristics of autistic brain cortical maps. A SOM can represent input features just like a brain map retains salient perceptual stimuli, and can exhibit comparable properties to an autistic brain map if its formation mechanism is impaired. The modeling premise of the impairment is suggested not by the biological map, but by its model. Specifically, Gustafsson argued that a biologically plausible cause of impairment in a SOM is the excessive lateral feedback inhibitory synaptic strengths that can degrade a map's generalization and feature representation capacity, resulting in high sensory discrimination and feature specificity, even to the point of instability.

3 The Autistic SOM Model

3.1 IPSOM: A Spatial Model

IPSOM (Interlocking Puzzle SOM) is a prototype SOM spatial behavioral model of how humans complete interlocking (jigsaw) puzzles [15]. The mathematical and algorithmic form of the neural network employed is according to Haykin [16]. When trained, using a representative sample of puzzle completion sessions, it forms a behavioral SOM of the statistically dominant patterns (strategies) of puzzle completion. IPSOM has been evaluated for the case of 4x5 (20-piece) puzzles against a simulated group of virtual people. Each virtual person was assumed to use one of four predetermined puzzle completion strategies (Fig. 1). The design principles behind the selected strategies were the generation of a small number of straightforward, real-life based

patterns, the utilization of topological clustering, and the prioritization of the basic strategy of determining the board periphery during the puzzle completion. IPSOM was conclusively found to be efficient in modeling the behavioral domain [15].

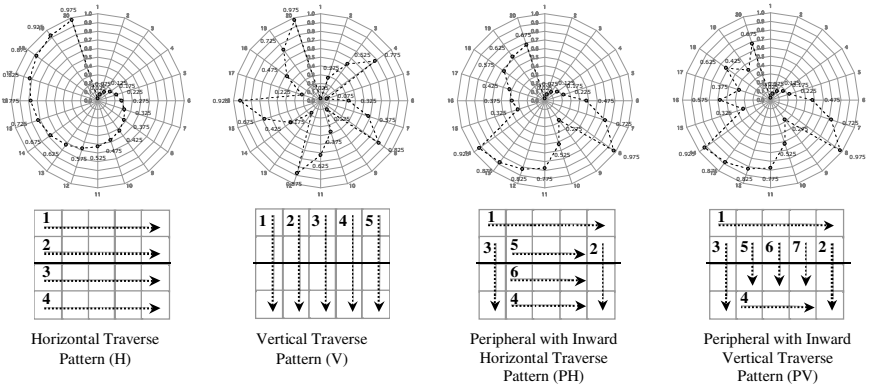

Fig. 1. An illustration of the four puzzle completion strategies used to evaluate IPSOM. A radar-graph depicts the order of puzzle completion for each pattern (H, V, PH, PV). The radial axis shows the encoded numerical position values on the puzzle board (i.e., which puzzle piece), and the angular axis shows the discrete completion sequence numbers (i.e., which piece is first, second, etc.) By connecting the points on the graph, a distinct visual pattern is formed. Attached to each graph, a puzzle board contains the puzzle completion order conventionally.

In this paper, IPSOM is employed as a modeling test-bed for cortical map spatial perception. The working hypothesis is that IPSOM not only is a behavioral model but also a cognitive model of how humans perceive puzzle completion strategies when presented with puzzle completion examples. It is assumed that an average person would form an internal representation of the dominant strategies; a cortical map would retain the domain specific knowledge, modeled by a trained SOM. The IPSOM map is expected to effectively depict the training patterns in a topologically ordered fashion, where neighboring patterns are also visually similar (Fig. 2).

	1	2	3	4	5	6
1	**PV**	PV	PV-	V-	V	**V**
2	PV	PV-	PV-	V-	V-	V
3	PV-	PV-	~	~	V-	V-
4	PH-	PH-	~	~	H-	H-
5	PH	PH-	PH-	H-	H-	H-
6	**PH**	PH	PH-	H-	H	**H**

X	Primary core neuron:	Optimal pattern representation
X	Core neuron:	Good pattern representation
X-	Weak neuron:	Poor pattern representation
~	Undecided neuron:	Transitionally excessive pattern

Fig. 2. An abstract illustration of a possible IPSOM 6x6 cognitive behavioral map, after being trained using the predesigned four-pattern set. Each SOM neuron is best-matched to one of the data set patterns (H, V, PH, PV) with a corresponding pattern-representation strength.

3.2 The Autistic IPSOM

Gustafsson's theory of autism [10] postulates that autistic cortical maps are inadequate or even undeveloped, and suggests the excessive lateral feedback inhibitory synaptic strengths as the best candidate causal factor. In a SOM, this can be expressed as a premature narrowing of the topological neighborhood (TN) during training; TN can be regarded as the "source of power" [17] in the autistic cognitive model. We maintain that the initial width of the TN function affects the map's representational capacity in a way directly applied to Gustafsson's ideas. By using a modified version of IPSOM, we perform an evaluation with a complex weight-encoding model. A non-autistic brain is expected to successfully represent all the dominant puzzle completion strategies. This can be modeled using IPSOM in its original parameter configuration.

A series of simulations have been performed with the initial width of the TN function set to a typical value of 3 (i.e., equal to the network's radius, as suggested by Haykin [16]). As a representative example, we visualized the last row of a resulting cognitive behavioral map that closely matches the topological configuration of Fig. 2, and calculated the Euclidean distance of Pattern H to all neurons in the map. Here, a smooth transition between patterns is apparent (Fig. 3a and 4a) and illustrates the map's ability to generalize without losing its capacity to accurately represent all statistically important features (i.e., puzzle completion strategies) from the input space (i.e., perceptual stimuli). Thus, feature memory is enabled, in which a subsequent perceptual stimuli session of a slightly different puzzle completion strategy can be associated with one of the existing patterns already stored in the map without significant matching errors. Transitional neurons are not only present but they also meaningfully indicate the map's perceptual stability. The resulting map is non-autistic.

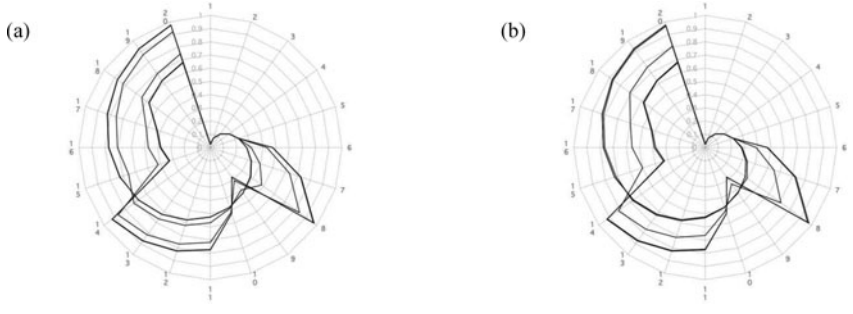

Fig. 3. Combined illustration of IPSOM map's last-row (six) neurons after two separate simulations (initial η_0=.1) for different initial TN widths (all other parameters identical). In (a), an initial width of 3 facilitates the representation of transition between patterns. In (b), a smaller width of 1.1 results in neurons tightly grouped into two patterns with impaired transition.

In a second series of simulations, the initial width of the TN function was narrowed to 1.1. Again, as a representative example, we visualized the last row of the resulting cognitive behavioral map that is directly comparable to the previous map discussed (i.e., using the same pseudorandom seed to randomly generate the network's initial

synaptic weights), and calculated the Euclidean distance of Pattern H to all neurons in the map. Here, it is evident that there is high pattern discrimination with poor transition between patterns (Fig. 3b and 4b). The map's topological ordering is largely reduced to jumps from pattern to pattern and to hit-or-miss representations indicating an inability to generalize from the input space. The feature memory becomes impaired; a subsequent perceptual stimuli session of a slightly different puzzle completion strategy cannot be associated with one of the existing patterns in the map without significant matching errors, if at all. Evidently, the resulting map exhibits autistic traits.

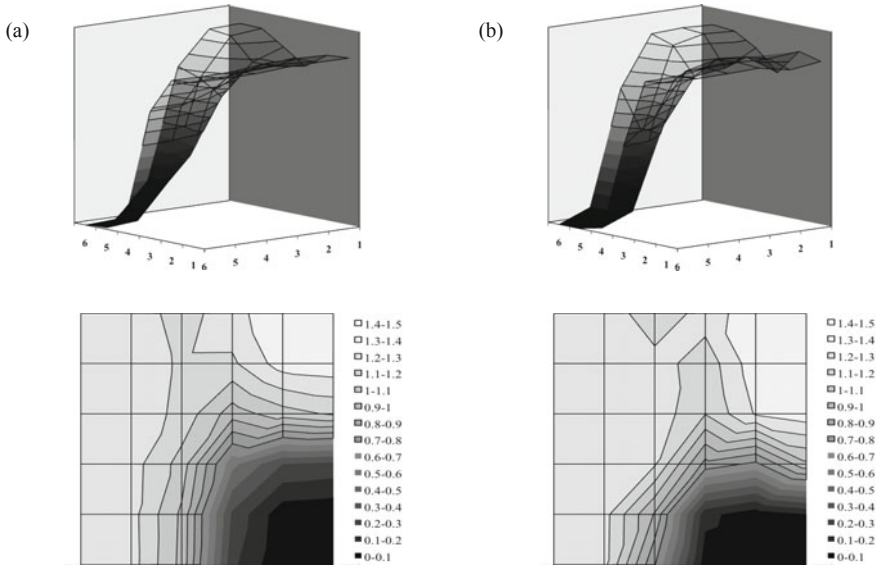

Fig. 4. The Euclidean distance of Pattern H to each neuron in the map is depicted on 3D and 2D graphs after two simulations (initial learning rate η_0=.1) for different initial TN widths (all other parameters identical). The darker and closer to the horizontal 3D base-plane (the IPSOM map) areas signify smaller distance and, thus, higher representational accuracy of Pattern H. In (a), an initial width of 3 facilitates a smoother transition from Pattern H to other patterns on the map, whereas in (b) a width of 1.1 results in steeper increase of the Euclidean distance indicating transitional impairment.

4 Implementation Details and Statistical Analysis

4.1 Technical Aspects of the Model

The IPSOM model is constructed in ANSI C. Its SOM is implemented in a three-dimensional lattice containing the synaptic weight vectors of the neural network. The first two dimensions represent the coordinates needed to refer to a specific neuron in the lattice, while the third dimension is used for holding the synaptic weight vector of that neuron. The training set used exhibited no bias towards any of the H, V, PH, and PV patterns. Each training was concluded in 950 epochs, with an initial learning rate of

.1, and consisted of lattice initialization, neuron competition, neuron cooperation, and synaptic adaptation. The TN of the winning neuron is specified by a translation invariant Gaussian function with exponential decay and it is based on the lateral distance between the winning neuron and the excited neuron. IPSOM uses stochastic approximation, which is implemented by a time-varying learning rate with exponential decay. Again, since the 'standard' SOM network was followed, Haykin [16] provides details on the mathematical formulas used. Each resulting map consists of 36 topologically ordered neurons, each of which holds a synaptic weight vector containing a pattern with varying degree of similarity to the four training patterns as discussed.

4.2 Statistical Analysis

Each pattern is encoded as a vector of 20 numbers representing a puzzle completion strategy; each number represents a specific puzzle piece. Depending on each piece's numerical value and order in the vector, it is possible to determine which puzzle piece was placed on the board during the puzzle completion task, and when.

In order to assess the strength of the association between the puzzle completion patterns of the training set, bivariate non-parametric correlation analysis was performed for each of them against the others. The analysis revealed that there is a highly significant positive association between the PH and PV patterns (ρ=.974, P<.001, N=20). Reliably positive associations were also found between the H and PH patterns (ρ=.523, P=.018, N=20), and between the H and PV patterns (ρ=.507, P=.023, N=20). The results indicate a very strong strategy-wise similarity between the PH and PV puzzle completion methods, and a strong strategy-wise similarity between the H and PH patterns. They also explain the partial visual compatibility observed between the H and PH patterns and the H and PV patterns. These results are in agreement with the patterns topology of a typical IPSOM (see also Fig. 2).

Since the IPSOM neurons have identical structure with the training patterns, bivariate linkage analysis was performed again as an exploratory examination method of the non-autistic and autistic SOM variants of IPSOM. For each SOM, the correlation coefficients of all the horizontally neighboring (immediate and more distant) neurons were computed. Table 1 presents results based on the last row of neurons of each IPSOM variant; the contents of these neurons are visualized in Fig. 3.

The results show that the pattern transition between the neurons in the non-autistic model (N1…N6) is smoother than in the autistic model (A1…A6) (see also Fig. 3). Specifically, the pattern transition from PH to H is completed in two neurons (N3, N4) in the non-autistic map, whereas it only takes a 'weak' PH neuron (A3) in the autistic map. The characteristically very high pattern discrimination in autistic brain cortical maps is evident in this case. In the last row of the autistic map, the neurons practically represent two tightly grouped patterns; the representational variation within each group (PH and H) is almost non-existent. Notably, in the selected case for this analysis, the correlation between any two neurons is by default statistically significant since the represented pattern is either a PH or an H, and a reliably positive association between the latter two has already been mentioned above. However, the degree of this positive association varies, as described, in exactly the way that would indicate an autistic phenotype.

Table 1. Correlation analysis of the non-autistic (N) and autistic (A) variants of IPSOM. The neurons N/A1…N/A6 correspond to non-autistic/autistic IPSOM neurons with map coordinates (6,1), …, (6,6), respectively, which are all the neurons of the last row on each map.

Spearman's ρ		*N1*	*N2*	*N3*	*N4*	*N5*	*N6*		*A1*	*A2*	*A3*	*A4*	*A5*	*A6*
Cor. Coef.	*N1*	1.0	1.0#	.979#	.595#	.523*	.523*	*A1*	1.0	1.0#	.896#	.523*	.523*	.523*
Sig.		.	.	.000	.006	.018	.018		.	.	.000	.018	.018	.018
Cor. Coef.	*N2*	1.0#	1.0	.979#	.595#	.523*	.523*	*A2*	1.0#	1.0	.896#	.523*	.523*	.523*
Sig.		.	.	.000	.006	.018	.018		.	.	.000	.018	.018	.018
Cor. Coef.	*N3*	.979#	.979#	1.0	.714#	.642#	.642#	*A3*	.896#	.896#	1.0	.797#	.797#	.797#
Sig.		.000	.000	.	.000	.002	.002		.000	.000	.	.000	.000	.000
Cor. Coef.	*N4*	.595#	.595#	.714#	1.0	.991#	.991#	*A4*	.523*	.523*	.797#	1.0	1.0#	1.0#
Sig.		.006	.006	.000	.	.000	.000		.018	.018	.000	.	.000	.000
Cor. Coef.	*N5*	.523*	.523*	.642#	.991#	1.0	1.0#	*A5*	.523*	.523*	.797#	1.0#	1.0	1.0#
Sig.		.018	.018	.002	.000	.	.		.018	.018	.000	.	.	.
Cor. Coef.	*N6*	.523*	.523*	.642#	.991#	1.0#	1.0	*A6*	.523*	.523*	.797#	1.0#	1.0#	1.0
Sig.		.018	.018	.002	.000	.	.		.018	.018	.000	.	.	.
N(pairwise)=20; Sig. (2-tailed): Correlation is significant at the 0.05 level (*) and at the 0.01 level (#).														

5 Conclusion and Future Work

It is reported in the literature that SOM neural networks resist formal analysis [18]. Nevertheless, we have been able to show that they can be very efficient in modeling even hard problems, like facets of autism, and can offer valuable scientific insights past the behavioral level. This study is part of an ongoing effort to provide a computational account of how autistic behavior is associated with specific functional and structural characteristics of the human nervous system, and, in particular, to investigate Gustafsson's claims about the role of inhibitory lateral feedback synaptic connection strengths in autism [10]. Analysis of the output of IPSOM, a novel and weight-encoding complex model, has provided evidence to support the latter.

Additional research is under way to investigate alternative formulations of the TN function with higher biological plausibility for a number of brain disorders including autism and delusions in schizophrenia [19]. In the course of this research, empirical examinations on impaired brain cortical maps of humans situated in controlled environments, as well as neurocomputational investigations of existing neuroscientific clinical data, will hopefully aid in revealing further critical modeling parameters.

References

1. Sun, R., Coward, L.A., Zenzen, M.J.: On Levels of Cognitive Modeling. Philosophical Psychology 18, 613–637 (2005)
2. Thomas, M.S.C., Karmiloff-Smith, A.: Connectionist Models of Cognitive Development, Atypical Development and Individual Differences. In: Sternberg, R.J., Lautrey, J., Lubart, T. (eds.) Models of Intelligence: International Perspectives, vol. 44, pp. 133–150. American Psychological Association, Washington, DC (2003)

3. Munakata, Y., McClelland, J.L.: Connectionist Models of Development. Developmental Science 6, 413–429 (2003)
4. Kanner, L.: Autistic Disturbances of Affective Contact. Nervous Child 2, 217–250 (1943)
5. Asperger, H.: Die Autistischen Psychopathen im Kindesalter. Archiv für Psychiatrie und Nervenkrankheiten 117, 76–136 (1944)
6. Gillberg, C., Coleman, M.: The Biology of the Autistic Syndromes. Cambridge University Press, Cambridge (2000)
7. Hermelin, B.: Images and Language. In: Rutter, M., Schopler, E. (eds.) Autism: A Reappraisal of Concept and Treatment. Plenum, New York (1978)
8. Happe, F.: The Autobiographical Writings of Three Asperger Syndrome Adults: Problems of Interpretation and Implications for Theory. In: Frith, U. (ed.) Autism and Asperger Syndrome. Cambridge University Press, Cambridge (1991)
9. Shulman, C., Yirmiya, N., Greenbaum, C.W.: From Categorization to Classification: A Comparison Among Individuals with Autism, Mental Retardation, and Normal Development. Journal of Abnormal Psychology 104, 601–609 (1995)
10. Gustafsson, L.: Inadequate Cortical Feature Maps: A Neural Circuit Theory of Autism. Biological Psychiatry 42, 1138–1147 (1997)
11. Frith, U.: Autism: Explaining the Enigma. Blackwell Publishers, Oxford (1989)
12. Hermelin, B., O'Connor, N.: Psychological Experiments with Autistic Children. Pergammon Press, Oxford (1970)
13. Willshaw, D.J., vor der Malsburg, C.: How Patterned Neural Connections Can Be Set Up by Self-Organization. In: Proceedings of the Royal Society of London, Series B, vol. 194, pp. 431–445. Royal Society, London (1976)
14. Kohonen, T.: Self-Organized Formation of Topologically Correct Feature Maps. Biological Cybernetics 43, 59–69 (1982)
15. Revithis, S., Wilson, W.H., Marcus, N.: IPSOM: A Self-organizing Map Spatial Model of How Humans Complete Interlocking Puzzles. In: Sattar, A., Kang, B.H. (eds.) AI 2006. LNCS (LNAI), vol. 4304, pp. 285–294. Springer, Heidelberg (2006)
16. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice-Hall, NJ (1999)
17. Sun, R., Ling, C.: Computational Cognitive Modeling, the Source of Power and Other Related Issues. AI Magazine 19, 113–120 (1997)
18. Cottrell, M., Fort, J.C., Pagès, G.: Theoretical Aspects of the SOM Algorithm. Neurocomputing 21, 119–138 (1998)
19. Revithis, S.: Significance of Topological Neighborhood in SOM Cognitive Modeling of Brain Disorders: Current Neurocomputational Simulations. Book of Abstracts of the 16th International Conference of the Association of Psychology & Psychiatry for Adults & Children - APPAC Journal, vol. 18(2), p. 26. APPAC, Athens (2011)