Title
SOM Cognitive Modeling of Autistic and Schizophrenic Traits Using an Oscillating Topological Neighborhood Width Function

Permalink
https://escholarship.org/uc/item/6963062m

Journal

ISSN
1069-7977

Authors
Revithis, Spyridon
Wilson, William
Marcus, Nadine

Publication Date
2013

Peer reviewed
SOM Cognitive Modeling of Autistic and Schizophrenic Traits
Using an Oscillating Topological Neighborhood Width Function

Spyridon Revithis (revithiss@cse.unsw.edu.au)
William H. Wilson (billw@cse.unsw.edu.au)
Nadine Marcus (nadinem@cse.unsw.edu.au)

School of Computer Science and Engineering, University of New South Wales
UNSW-Sydney, NSW 2052, Australia

Abstract
The artificial neural network class of self-organizing maps (SOMs) is a powerful and promising cognitive modeling tool in the study of the brain and its disorders. Under this premise, this paper proposes a novel modification of the standard SOM algorithm in the form of an oscillating Topological Neighborhood (TN) width function. Existing research in neuroscience indicates that SOMs with oscillating TN width could exhibit higher biological plausibility than standard TN width SOMs. In this paper, two neuro-developmental disorders, autism and schizophrenia, are modeled, based on existing neurocomputational theories, using both SOM approaches. The simulation results demonstrate that there is significant equivalence between standard and oscillating TN width SOM modeling in terms of map formation behavior, output and structure. The theoretical and computational arguments presented validate the proposed SOM modification within a cognitive modeling framework.

Keywords: Self-Organizing Maps, Cognitive Modeling, Cortical Maps, Autism, Delusions, Schizophrenia.

Introduction
Computational modeling offers a powerful way to study cognition and behavior. It has been applied to numerous areas of psychology and provides a more promising framework than those based on verbal models in terms of methodological diversity and applicability potential (Sun, Coward & Zenzen, 2005). An ever-increasing number of computational modeling studies are dedicated to the modeling of cognitive and developmental phenomena using artificial neural networks (Thomas & Karmiloff-Smith, 2003; Polk & Seifert, 2002; Parks, Levine & Long, 1998).

Shultz (2003) provides a comparative evaluation of the different computational neural network systems used to model cognitive developmental phenomena. An important class of such modeling networks is the self-organizing feature map; it is based on a Hebbian-type (Hebb, 1949) unsupervised neural learning mechanism and uniquely resembles topographic cortical maps in the brain to which has directly comparable structure and output characteristics (Spitzer, 1995b; Livingstone & Hubel, 1988; Blasdel & Salama, 1986; Merzenich & Kaas, 1980). Willshaw and von der Malsburg (1976) originally proposed the self-organizing neural network to account for the retinotopic mapping problem. Kohonen’s version (2001) -commonly abbreviated to ‘SOM’-, however, possesses significant computational characteristics and a range of powerful properties, particularly relevant to understanding and modeling of cortical brain maps, including approximation of the input space, topological ordering, density matching, and feature selection (Haykin, 1999).

This study investigates cognitive modeling aspects of modeling neuro-developmental disorders using SOM neural networks. The first section presents the SOM modeling framework used in this work, and introduces a novel modification in the SOM formation algorithm with significant cognitive modeling implications. In the subsequent two sections, core biological and behavioral characteristics of two mental disorders, autism and schizophrenia, respectively, are modeled using a prototype SOM model. The last section provides a discussion of the computational and theoretical parameters of the SOM modeling employed in the paper.

The SOM Modeling Framework

Aspects of SOM Neural Networks
A SOM is a non-linear unsupervised-learning computational neural network consisting of two layers. It has the capacity to map an input ‘environmental’ layer, consisting of patterns of fixed but arbitrary dimension, onto a (usually) one or two dimensional lattice ‘representational’ layer. The representation of environmental input in the output layer (called the map) is performed in a topologically ordered fashion, maintaining the non-linear input data distribution, and involves dimensionality reduction. Figure 1 shows an abstract depiction of a two-dimensional SOM; each input layer pattern vector connects fully with the map neurons.

![Figure 1: A two-dimensional SOM.](image-url)
The SOM neural network formation (training) process has four parts (as described in Haykin (1999)): synaptic weight initialization of the output lattice; neuron competition; neuron cooperation; and synaptic adaptation. The last three are sequenced within a loop for a finite number of ‘epochs’, in which input patterns are presented and weights adjusted until the weights converge.

During the competition phase, a winning neuron for the current input pattern is determined, based on a Euclidean distance metric. In the cooperation phase the winning neuron becomes the center of a cooperative process extending around an area according to a topological neighborhood (TN) function. In the synaptic adaptation phase, the weights of the map neurons within the TN of the winning neuron are updated ‘towards’ the current input pattern at an intensity determined by their lateral distance to the winning neuron as well as an exponentially decaying learning rate function.

From a cognitive modeling perspective, it is of particular interest to examine the neurobiological relevance of the SOM formation process at the implementation level of the neuron lateral interaction and inhibition mechanism. The standard SOM algorithm (Haykin, 1999) employs a translation invariant Gaussian TN function with an exponentially decreasing width, as illustrated in Figure 2.

![Figure 2: Decreasing TN width around a winning neuron (dark grey neuron) in a two-dimensional SOM.](image)

The TN width function can be expressed by the formula

\[ \sigma(n) = \sigma_0 \cdot \exp\left(-\frac{n}{\tau_1}\right), \quad n = 1, 2, \ldots, t \]

where \(\sigma_0\) is the initial TN width, \(\tau_1\) is a time constant, \(t\) is the number of epochs, and \(n\) is the current epoch.

The fact that only neurons close to the winning neuron have their weights changed significantly (implemented at the biological neural network level by a mixture of excitation and lateral inhibition) has a measurable impact on the representational structure of the SOM. A number of SOM cognitive models of brain disorders center around the key role of TN width and its exegetic biological significance (Gustafsson, 1997; Spitzer, 1999).

**Oscillating TN width SOM**

The SOM cooperative phase involves local neuronal interactions via group Hebbian activation regulated by lateral inhibition. In general, neural synchrony and communication at the local and long-range level is an important aspect of brain functioning; neural oscillation, particularly correlated to inhibitory neural activity, is increasingly considered to be of paramount importance to neural information processing and central to a number of studies of mental disorders including schizophrenia and autism (Schnitzler & Gross, 2005; Wang, 2010). Neuronal group oscillatory synchrony is linked to inhibitory interneuron rhythmic modulation of the firing rate of excitatory neurons, at the local interaction neuronal level (Cardin, Carlen, Meletis, Knoblich, Zhang, Deisseroth, Tsai & Moore 2009). Last, synchronous oscillatory activity of neighboring inhibitory interneurons may be supported by sub-threshold oscillatory behavior (Llinas, 1988).

In line with the relevant research on neural oscillation outlined above, this paper introduces a modification with increased biological plausibility in the SOM cooperative phase, as previously reported in a preliminary study (Revithis, 2011). Specifically, the original TN width function, part of the overall TN function, is replaced by a new TN width function that exhibits local exponential decrease instead of global. In this way the TN width oscillates continuously throughout the SOM formation process. Oscillation is necessary in a biologically plausible model, otherwise learning would cease when the TN approached zero. The oscillation consists of a concatenation of exponentially decreasing original TN width ‘function instances’; thus, in the same number of epochs (i.e., one SOM training session) multiple function instances will fit, as shown in Figure 3.

![Figure 3: SOM oscillating TN width.](image)

The new function can be expressed as

\[ \sigma'(n) = \sigma_0 \cdot \exp\left(-\frac{(n+1)\mod t'}{\tau_1'}\right), \quad n = 0, 1, 2, \ldots, t-1 \]

where \(\sigma_0\) is the initial TN width, \(\tau_1'\) is a time constant, and \(n\) is the current epoch. The constant \(t' = t / c\), where \(c\) is the oscillation constant determining how many times the TN width will reset to \(\sigma_0\) and start decreasing again.

**IPSOM**

IPSOM (Interlocking Puzzle SOM) is a complex-weight-encoding prototype SOM spatial behavioral model of how humans complete interlocking puzzles (Revithis, Wilson & Marcus, 2006). When trained, using a representative sample of puzzle completion sessions, it forms a behavioral SOM of the statistically dominant patterns (strategies) of puzzle completion. A 6x6 IPSOM has been evaluated for the case of 4x5 puzzles against a simulated group of people. Each ‘virtual’ person used one of four predetermined puzzle completion strategies, illustrated in Figure 4.

![Figure 4: IPSOM strategy examples.](image)
Each radar-graph in Figure 4 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. 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 emphasizing the basic strategy of determining the board periphery during the puzzle completion. IPSOM was conclusively found to be efficient in modeling the behavioral domain.

**Figure 4:** IPSOM training set patterns (strategies).

In this paper, IPSOM is employed as a modeling test-bed for cortical map spatial perception. The working hypothesis is that IPSOM is not only 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 representation of the dominant strategies; a cortical map spatial perception. The working hypothesis in modeling the behavioral domain 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. 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 emphasizing the basic strategy of determining the board periphery during the puzzle completion. IPSOM was conclusively found to be efficient in modeling the behavioral domain.

**Figure 4:** IPSOM training set patterns (strategies).

**Modeling Aspects of Autism using IPSOM**

**A Neural Circuit Theory of Autism**

Autism, a pervasive developmental disorder, has been studied for over 50 years by an expanding interdisciplinary research community. The current diagnostic tools (DSM-IV and ICD-10) dictate a socio-psychological behavioral approach that does not inform of the causes of autism; nevertheless, it is considered to be neurobiological in nature (Coleman & Gillberg 2012).

Autism is associated with atypical perception and its internal representation. Sensory input often fails to integrate into existing memory due to abstraction impairment; there is difficulty in detecting the important features among the non-essential details; elaborating on internal representations is also problematic, where it appears that central executive control is required (Frith, 2003).

Gustafsson’s (1997) neural circuit theory of autism 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 (Frith, 2003). Neurological deficiencies in the formation of brain cortical maps give rise to autistic attributes. 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 (Gustafsson, 1997). Autistic maps lack feature distinction and preservation, and fail to provide an internal representation of salient perceptual data leading to raw data memory that lacks sophisticated representations.

According to Gustafsson (1997), SOMs provide a biologically plausible way to model characteristics of ‘autistic’ cortical maps. A SOM can represent input features just as a cortical map in the brain retains salient perceptual stimuli, and can exhibit similar deficiencies to an autistic cortical map if its formation mechanism is impaired.

**The Autistic IPSOM**

The modeling premise of the SOM autistic impairment is suggested not by the biological map, but by its model. Gustafsson (1997) argued that a biologically plausible cause of impairment in a SOM is the application of excessive lateral feedback inhibitory synaptic strengths. The latter can degrade the map’s generalization and feature representation capacity, resulting in high sensory discrimination and feature specificity, even to the point of instability, leading to the formation of inadequate or even undeveloped maps.

This modeling premise can be expressed as a TN premature narrowing during SOM training; TN can be regarded as the “source of power” (Sun & Ling, 1997) in the autistic model. The initial TN width ($\sigma_0$) in the TN width function affects the map’s representational capacity in a directly applicable way to Gustafsson’s theory (Revithis & Tagalakis, 2012). A non-autistic cortical map is expected to
represent all the dominant puzzle completion strategies with smooth transition between them. This can be modeled using IPSOM in its original parameter configuration.

After the incorporation of TN parameter modifications on IPSOM, an evaluation was performed. A series of groups of controlled simulations were executed with the initial width of the TN function set to a typical value of \( \sigma_0 = 3 \) (i.e., equal to the network’s radius, as suggested by Haykin (1999)) for one group, and reduced to \( \sigma_0 = 1.15 \) for another group. Both groups were executed twice, using a standard TN width function, in one simulation series, and an oscillating TN width function in a second one. The results (discussed next) from over 150 simulations confirm that, for large \( \sigma_0 \), the resulting IPSOM exhibits efficient representation of the input space, whereas IPSOM training, using a small \( \sigma_0 \), forms a map with autistic structural characteristics. The results also support the hypothesis that the oscillating TN width IPSOM is equivalent to the standard TN width IPSOM in modeling autistic traits.

![Figure 6: Standard TN width IPSOM map characteristics.](image)

Figure 6 depicts IPSOM neurons after training, using a standard TN width function, for \( \sigma_0 = 3 \) (top) and \( \sigma_0 = 1.15 \) (bottom). The leftmost 3D graphs, and the 2D graphs in the middle, depict the Euclidean distance of pattern H to each neuron in the map. The darker and closer to the horizontal 3D base-plane (map) areas signify smaller distance and, thus, higher representational accuracy for pattern H. A \( \sigma_0 = 3 \) facilitates a smoother transition from pattern H to other patterns in the map, whereas a \( \sigma_0 = 1.15 \) results in steeper increase of the Euclidean distance indicating transitional pattern impairment. The rightmost combined-concentric radar graphs depict five neighboring IPSOM neurons for \( \sigma_0 = 3 \) (top) and \( \sigma_0 = 1.15 \) (bottom). A \( \sigma_0 = 3 \) facilitates smoother transition from Pattern H to V, whereas for \( \sigma_0 = 1.15 \) neurons are tightly grouped in two patterns (H and V) with impaired transition and generalization capacity.

Figure 7 depicts IPSOM neurons after training, using an oscillating TN width function, for \( \sigma_0 = 3 \) (top) and \( \sigma_0 = 1.15 \) (bottom). The observations that can be made are identical to the ones of Figure 6.

![Figure 7: Oscillating TN width IPSOM map characteristics.](image)

The illustrated example-simulation-results of Figures 6 and 7 are representative of the totality of simulation results obtained in terms of the observed characteristics. Patterns H and V, which were used for the rightmost concentric radar graphs, were selected to better demonstrate IPSOM’s transitional behavior due to their relatively low correlation significance amongst IPSOM training set patterns (Table 1).

<table>
<thead>
<tr>
<th>Spearman’s ( \rho ) (N=20)</th>
<th>H</th>
<th>V</th>
<th>PH</th>
<th>PV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>H</td>
<td>1</td>
<td>.429</td>
<td>.523</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td>.059</td>
<td>.018</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>V</td>
<td>1</td>
<td>.388</td>
<td>.420</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td>.091</td>
<td>.066</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>PH</td>
<td></td>
<td>1</td>
<td>.974a</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>PV</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Correlation is significant at the 0.05 level (*) and at the 0.01 level (#).

### Using IPSOM to Model Delusions

#### Acute and Chronic Delusions in Schizophrenia

Modern studies on schizophrenia span approximately a century. There has been a continuous evolution of the understanding of this mental disorder and currently it is widely considered to be a progressive neuro-developmental disorder. Amongst its common positive psychotic symptoms are delusions (Green, 2001).

Spitzer has argued (1995a, 1995b, 1999) that SOM neural networks can provide a model of brain cortical function, and implement lateral inhibition, an essential feature of cortical maps. Furthermore, he proposed a neurocomputational exegetical framework for delusions based on the concepts of neuromodulation and neuroplasticity in relation to formation and operation of sensory and higher-order computational maps in the cortex.

Specifically, according to this approach, neuromodulator activity in the brain is associated with the signal-to-noise
ratio at the neuronal level, from an information-theoretic perspective. High neuromodulator activity can lead to an increase of focusing in neuronal activation and is associated with acute delusional states; such focusing can be modeled via excessive SOM lateral inhibition. Chronic delusions can then be regarded as the result of the establishment of entrenched cortical maps via sustained acute delusional states due to brain neuroplasticity.

**IPSOM Modeling of Delusions**

According to Spitzer (1995a), a decisive factor in the clinical phenomenon of acute delusions is the level of cortical neuromodulator activity; this affects modulation of signal-to-noise ratio. In a SOM model of delusions it is possible to regulate the level of neuronal activation focusing associated with the signal-to-noise ratio by controlling SOM lateral inhibition. This can be achieved by controlling the width of TN during SOM formation. TN can be regarded as the “source of power” (Sun & Ling, 1997) in this model.

Similar to the autistic model, the working hypothesis is that the initial TN width (\(\sigma_0\)) in the TN width function affects the map’s behavior in a way applicable to Spitzer’s theory. Inducing acute delusions in IPSOM can be realized via modifying the cooperation phase of the SOM algorithm in the model to employ a significant TN narrowing.

A series of groups of controlled simulations were executed with the initial width of the TN function set to a typical value of \(\sigma_0=3\) for one group, and reduced to \(\sigma_0=1.15\) for another group, as in the autistic model. Both groups were executed twice, using a standard TN width function, in one simulation series, and an oscillating TN width function in a second one. The results (discussed next) from over 150 simulations confirm that, for large \(\sigma_0\), the resulting IPSOM exhibits typical representation of the input space; when a small \(\sigma_0\) is used, however, the map’s formation behavior is atypical and retains structures corresponding to chronic delusions. The results also support the hypothesis that the oscillating TN width IPSOM is equivalent to the standard TN width IPSOM in modeling delusions.

Entrenched SOM structures that could give rise to chronic delusions can be identified by comparing ‘suspected’ formed IPSOM maps with their untrained (initial) state. A ‘delusional’ structure can plausibly be seen as a number of trained neurons representing neither a transitional pattern nor an input space pattern, or, excessively representing an input space pattern (the latter can be regarded as compromising the SOM density matching property (Haykin, 1999)). Furthermore, representational resistance to change can also be interpreted as a characteristic of established (chronic) delusional structures (Spitzer, 1995a).

Figure 8 depicts four snapshot graphs of the same part of the IPSOM map for different initial parameters. In graph A we see the situation before training - essentially random patterns, and the remaining three depict the map’s area after training for different \(\sigma_0\) value and TN width function configurations. By comparing IPSOM’s untrained graph with its standard TN width trained counterpart (graph C) we immediately observe the perseverance of a number of initial ‘blank’ patterns. A number of IPSOM neurons represent either the original initial ‘blank’ pattern or a distorted version of it. In the oscillating TN width case (graph B) there is also an excessive representation of the V pattern (cf. Figure 4). The observed ‘delusional’ flags, especially the resistance to environmental change, are prominent in the IPSOM trained graph using a very small \(\sigma_0\) (graph D).

![Figure 8: Induced delusional structure on IPSOM.](image)

**Discussion**

The significance of TN in SOM cognitive modeling has theoretical and practical implications. In this paper, a modified TN width function with increased biological plausibility (paramount to modeling) was introduced and simulation results, based on the IPSOM prototype, on two models of neuro-developmental disorders were presented.

The modeling significance of the oscillating TN width function is associated not only with the initial TN width (\(\sigma_0\)) parameter but, primarily, with the TN width ‘area’ covered throughout the SOM training. What is considered ‘narrow’ or ‘wide’ TN during SOM formation is -from a different perspective- a function of the TN width area covered.

![Figure 9: Standard and Oscillating TN width areas.](image)
In Figure 9, both the standard and the oscillating TN width functions are overlaid in both graphs. The TN width area has as an upper bound the corresponding TN width function and as a lower bound the epoch (horizontal) axis. Mathematically, the TN width area is expressed as
\[
\sigma(x) \text{ area} = \int_{0}^{T} \sigma_{0} \cdot e^{-x^{2}} \, dx = \sigma_{0} \cdot (\tau) \cdot e^{-C} + \sigma_{0}, \quad \sigma_{0}, \tau \in R.
\]
To calculate the area for a given TN width function, \(\sigma_{0}\), and number of epochs \(t\), the following formula was used:
\[
\sigma(x) \text{ area} = \int_{0}^{\tau} \sigma_{0} \cdot e^{-x^{2}} \, dx
\]
In the standard \& oscillating TN width IPSOM simulation results, the calculated \(\sigma(x) \text{ area}\) (for the same \(\sigma_{0}\) remained unchanged irrespective of the TN width function used. This verifies the output equivalence between the two modeling approaches. Furthermore, when, in the oscillating TN width function simulations, the \(\sigma_{0}\) value was reduced to \(\sigma_{0}'\), the calculated \(\sigma(x) \text{ area}\) was significantly smaller (Figure 9, right graph) and resulted in an IPSOM map with more pronounced delusional structures (Figure 8, graph D). This demonstrates the computational and cognitive modeling significance of the TN width area.

In conclusion, it is important to note that making a link between the biological and computational levels, in such modeling studies, often requires a sequence of finely drawn associations across disparate disciplines. However indirect and interdisciplinary such a link may be, the methodology and tools to construct it have long been available, and an effort was made in this study to illustrate it.

References


