

Pattern Recognition in a Bucket

Chrisantha Fernando and Sampsa Sojakka

School of Cognitive and Computer Sciences , University of Sussex, BRIGHTON,
BN1 9QH, UK

{ctf20,sampsas}@cogs.susx.ac.uk

Abstract. This paper demonstrates that the waves produced on the surface of water can be used as the medium for a “Liquid State Machine” that pre-processes inputs so allowing a simple perceptron to solve the XOR problem and undertake speech recognition. Interference between waves allows non-linear parallel computation upon simultaneous sensory inputs. Temporal patterns of stimulation are converted to spatial patterns of water waves upon which a linear discrimination can be made. Whereas Wolfgang Maass’ Liquid State Machine requires fine tuning of the spiking neural network parameters, water has inherent self-organising properties such as strong local interactions, time-dependent spread of activation to distant areas, inherent stability to a wide variety of inputs, and high complexity. Water achieves this “for free”, and does so without the time-consuming computation required by realistic neural models. An analogy is made between water molecules and neurons in a recurrent neural network.

1 Introduction

Maass et al [1] have produced a model called the Liquid State Machine (LSM) which hypothesises that a cortical microcolumn consists of stereotyped recurrent circuits of integrate-and-fire neurons connected randomly according to a few parameters. The dynamics of this circuit are designed such that it can act as universal analog fading memory. Transient internal states can be read out by linear readout elements to produce stable outputs due to the high-dimensionality of the LSM. This allows robust real-time processing of time-varying inputs. All that needs to be trained is a readout module consisting of linear perceptrons.

The attraction of this model from an evolutionary robotic and a developmental neuroscience point of view is that it suggests a generic cortical architecture capable of universal real-time computation, which could potentially be specified by relatively few genetic parameters, so long as a “separation property” could be achieved¹. Further unsupervised learning would only need to take place on a layer of linear readout elements instead of on the recurrent neural network itself.

¹ The separation property requires that separation between the trajectories of internal states of the system is roughly proportional to the distances between two different input streams that caused them.

Here we have taken the metaphor seriously and demonstrated that real water can be used as an LSM for solving the XOR problem and undertaking speech recognition in the Hopfield and Brody “zero-one” discrimination [2]. Doing so is computationally efficient, since water undertakes the transformation from the low-dimensional input space of motors stimulating its surface into a higher dimensional space of waves in parallel. This is achieved by placing a glass bucket of water on top of an over-head projector, and filming the interference patterns which result from vibrating the surface of the water using motors as inputs, and feeding the resulting filmed wave patterns into a perceptron. This work relates to a physical simulation of a liquid surface produced by Goldenholz [6].

Our approach differs from recent work on real non-linear media for computing. For example, Andrew Adamatzky [4] [5] has solved the XOR problem using diffusive wave front interactions in two-reactant media in a T-maze, where in the [1 1] case, interference at the junction of the T-maze cancels out the reaction so producing the necessary non-linearity. In our bucket, it is not necessary to impose task-specific structural constraints that shape the wave fronts. We will show that the perceptron is able to discover the relevant information in a remarkably large variety of wave patterns. In addition, the media does not require a chemical reaction which for example produces a precipitate that reaches a point attractor, i.e. leaves a stable pattern of sediment in some stable state that represents the input pattern. In our system the perceptron must use only real-time wave dynamics to make a stable classification.

This work relates indirectly to work on field computation which acknowledges that neural computation is a massively parallel, low speed, low-precision continuous set of analog operations [7]. Serial computations must be short as dictated by the real-time response requirements of animals. MacLennan writes “The neural mass is sufficiently large that it is useful to treat it as a continuum”, and hypothesises that a similar architecture will be required to achieve artificial intelligence. He claims VLSI technology will not achieve this i.e. 15 million neurons/cm². The use of water-like media may serve as an alternative.

This mesoscopic approach is in contrast to mainstream neuroscience and neural networks which operates in a paradigm influenced by the work of McCulloch and Pitts [8], which views neurons as performing logical operations in circuits. Evidence for mesoscopic features relevant in conditioning tasks has been obtained by Walter Freeman [9] who describes spatial patterns of amplitude modulation in EEGs that change consistently following learning.

From an engineering perspective, the technological benefits of such an approach are within sight. Recent work by Ian Walmsley [10] in optics has demonstrated the power of exploiting interference patterns produced by the interaction of multiple sources of information to search a database consisting of 50 entries. Acoustic waves are used to encode information in the pattern of expansion and compression in tellurium dioxide. Coherent light of different frequencies (colours) is passed through the medium. Each bit of encoded information is probed by a different colour. Phase shifts due to the differences in thickness of the media (which encode the information) result in constructive or destructive interference

at the other end. By examining which frequencies have passed through in-phase, the information in the database can be retrieved.

Finally, we analyse the complexity of the water using a measure of complexity devised by Sporns, Tononi and Edelman [11] and show that water satisfies an approximation to the separation property described by Maass.

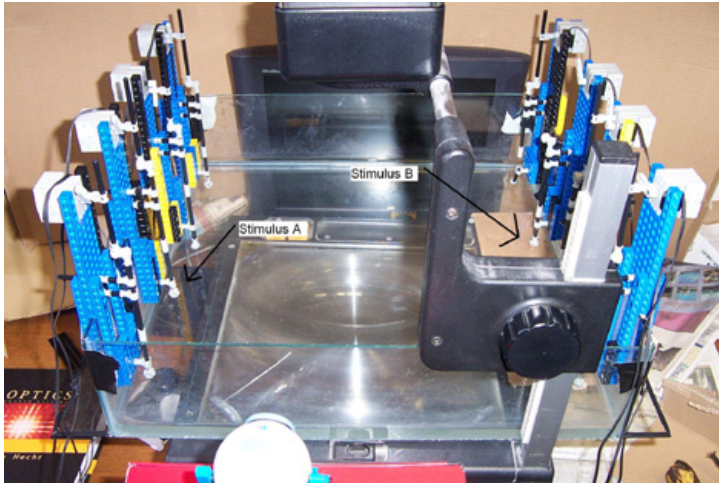


Fig. 1. The Liquid Brain.

2 Methods

The “liquid brain” was constructed from a transparent water tank suspended over an overhead projector, (Figure 1). Four electric motors connected to a computer via a serial control box were attached to each side of the tank to drive a small weight into the water and thus create eight wave point sources. The interference patterns so produced were projected on an antireflective cardboard and recoded at a resolution of 320x240 pixels at 5 frames per second using a standard web-cam. Each frame obtained with the camera was convolved using a Sobel edge detection mask that was suitably thresholded to remove noise and averaged to produce a 32x24 grid. The cells of this matrix were used as the input values for 50 perceptrons in parallel which were trained using the same p-delta rule as used by Maass et al [3].

2.1 XOR Task

In order to recreate the XOR problem using the liquid, two motors were set up at opposite sides of the tank. A single motor driven at constant frequency on left and right side of the tank corresponded to the $[1\ 0]$ and $[0\ 1]$ cases respectively. In the $[1\ 1]$ case both motors were activated simultaneously, and incoherently (i.e

with no fixed phase relation). One minute footage was recorded with the video camera of each of the four cases and the frames processed as discussed above yielding a total of 3400 frames. The order of the frames in this data set was randomised and used in training the linear readout perceptron. Another set of footage was taken of the four conditions as a test set.

2.2 Speech Recognition

The speech recognition task was designed to measure the suitability of the liquid brain model for robust spatiotemporal pattern recognition in a noisy environment. Earlier experiments have shown that transient synchrony of the action potentials of a group of spiking neurons can be used for “signal” recognition of a space-time pattern across the inputs of those neurons [2]. We show that the dynamical properties of the liquid can be used to produce this synchrony and that a simple readout neuron will suffice.

Following the example of the Hopfield and Brody experiments the speech recognition task used was to distinguish a speaker saying the word “one” from the same speaker saying “zero”. Twenty samples were recorded of a person saying each word with randomly varying amplitudes and intonations. The amplitude variation was mostly due to fluctuations in the distance of the speaker from the microphone ($\pm 10\text{cm}$). The voice samples were recorded as 12 kHz pulse-code modulated wave files ranging from 1.5 to 2 seconds in length.

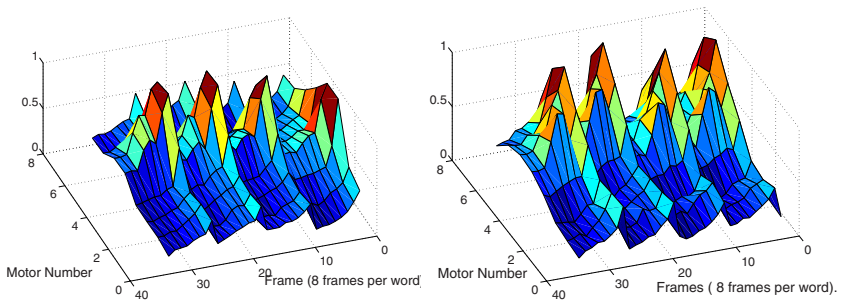


Fig. 2. Left: Four presentations of “Zero”. Right: Four presentations of “One”. Note the variability between and within sets.

Due to limitations in the speed of frequency response of the available motors a certain amount of pre-processing of the sound samples was necessary. Each sample was first processed using a Short-Time Fourier transform with a sliding window size set to the closest power of 2 of an eighth of the sample length thus splitting each sound sample to eight time slices. The active frequency range of 1-3000Hz of the resulting function of time and frequency was further averaged to eight frequency bands resulting in an 8x8 matrix for each sound sample, see Fig 2. The 40 matrices produced in this manner were first used to train a linear

perceptron by using the p-delta rule in order to ensure that the task was not linearly separable. Each sample matrix was then used to drive the motors of the liquid so that each of the eight frequency bands controlled one motor for 4 seconds (0.5 seconds per time slice). The 20 pre-processed samples of “one” were inserted into the liquid one after another and footage recorded of the entire sequence. This procedure was then repeated for the 20 samples of the word “zero”. The resulting frames were processed as above to create data sets used for the training of the readout perceptrons.

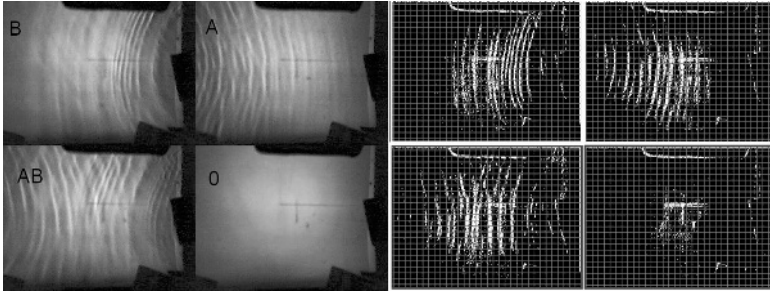


Fig. 3. Typical wave patterns for the XOR task. Top-Left: $[0\ 1]$ (right motor on), Top-Right: $[1\ 0]$ (left motor on), Bottom-Left: $[1\ 1]$ (both motors on), Bottom-Right: $[0\ 0]$ (still water). Sobel filtered and thresholded images on right.

3 Results

3.1 The XOR Problem

Figure 3 shows the waves that represent input conditions of the XOR task. Convergence to an optimal solution is obtained within 200 epochs (Figure 4) with no mistakes being made on the training set by this time. The summed outputs of the 50 perceptrons and their “correctness” for training and test sets is shown in Figure 5. The weight-matrices of these perceptrons show how the problem is being solved, see Figure 6. How can the perceptrons make the same classification in cases where the activities of inputs are radically different, i.e. both motors on versus still water? All the perceptrons learned a stereotyped weight matrix consisting of a central negative region and a lateral positive region of weights. Remarkably, this is a center-surround receptive field, with a vertical orientation bias. Thus, the solution does not in fact depend on interference patterns but rather upon mean amplitudes of wave fronts being different in different regions of the grid.

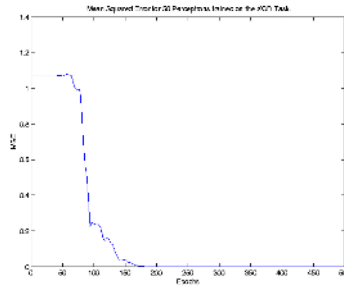


Fig. 4. Decreasing Mean Squared Error for 50 perceptrons trained on the XOR task after pre-processing by water.

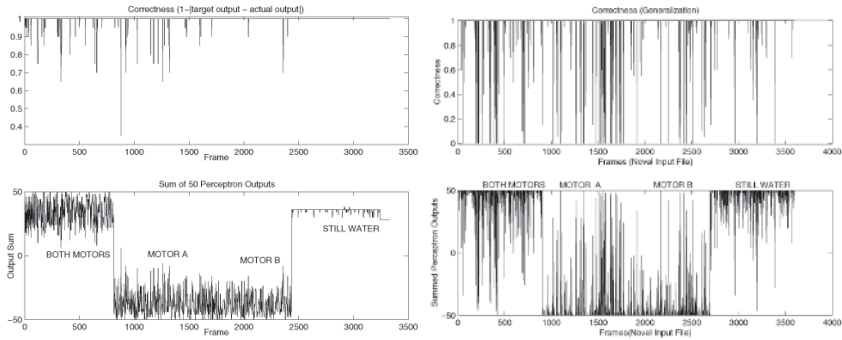


Fig. 5. Left : The summed outputs of the 50 perceptrons for the training data are shown in the bottom graph. Correctness in the top graph is defined as $(1 - \|\text{desiredoutput} - \text{actualoutput}\|)$ of the perceptrons after being passed through a stepwise squashing function. The readout module classifies the “still” and “both motor” conditions as class 1 and the “motor a” and “motor b” conditions as class -1. Although some perceptrons make mistakes, taking an average allows perfect classification. Right : Generalisation although not perfect, still results in 85% of frames being classified correctly.

3.2 Speech Recognition

The XOR problem, although non-linear is pretty easy to solve by adding just one extra dimension². Our approach of adding an extra 700 dimensions is to use a machete to garnish canapes. A significantly more difficult non-linear problem is speech recognition. Speech data when fed directly into a perceptron could not achieve a classification which made any fewer than 25% mistakes, However, error decreased to 1.5% when the input was passed through water, see Figure 7. Figure

² This can be done in various interesting ways. For example, imagine a robot with two arms. [0 0] the robot stands still, and does not fall over, [0 1] the robot pushes against the left wall with its left arm and falls over, [1 0] the robot pushes against the right wall with its right arm and falls over, [1 1] the robot pushes against both walls with both arms and so does not fall over. Inman Harvey, Personal Communication.

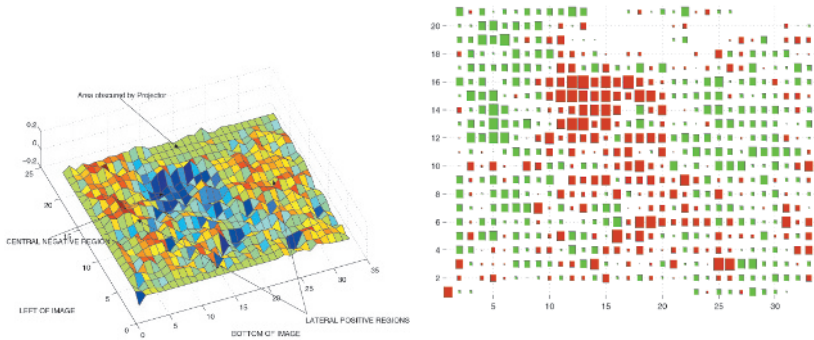


Fig. 6. The mean weight matrix of 50 perceptrons is shown as a surface plot on the left and as a Hinton diagram on the right. The center-surround receptive field works as follows. In the case of still water, positive and negative weights are activated uniformly and cancel each other out. In the case of both motors being stimulated, activity is also relatively uniform across the grid, and cancellation also occurs. However, with the $[1\ 0]$ or $[0\ 1]$ case, only the left or right of the field receives significant activation by wave fronts and the output of the perceptrons have a net positive value.

8 shows the summed outputs and correctness achieved by the 50 perceptrons on the training data and on the test set consisting of a novel set of footage of spoken words “zero” and “one”. Generalisation is relatively poor because of the noise. In our setup the frequency response of the motors differed widely and the performance of the shafts used to drive the water deteriorated with use, the water level in the tank and hence the amplitude of the waves fluctuated due to evaporation, vibrators moved out of place, the camera and the tank itself moved, the frame rate of the camera varied and no special attempts were used to remove these sources of noise. It is remarkable any classification could be made at all. We are confident that a more pristinely engineered version of the experimental setup would eliminate most of these problems, resulting in improved performance.

The mean weight matrix for these perceptrons consist of a set of wave front detectors, positioned at the top of the image, see Figure 9. This is the region with the most distinct wave fronts as the motors are closest to the filmed area of water.

4 Discussion

4.1 Measuring the Complexity of Water

Why does using water as a pre-processor work? Here we suggest that redundancy in the representation of information in the system may allow the perceptron to solve the problem with one of many patterns of weights. An information theory measure of neural complexity by Tononi, Sporns and Edelman has been used to show that brains are high in complexity due to a pattern of intrinsic connectivity whereby groups with similar response selectivities are preferentially connected,

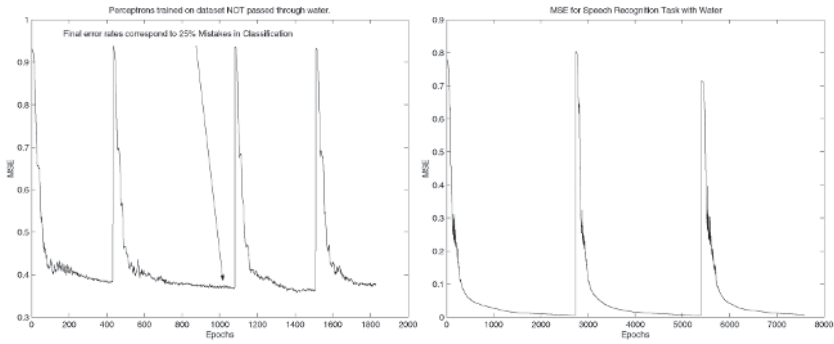


Fig. 7. Left: Error on raw data passed directly to perceptron. An Optimal classification is not possible. Right: Data passed through liquid. Optimal classification is possible.

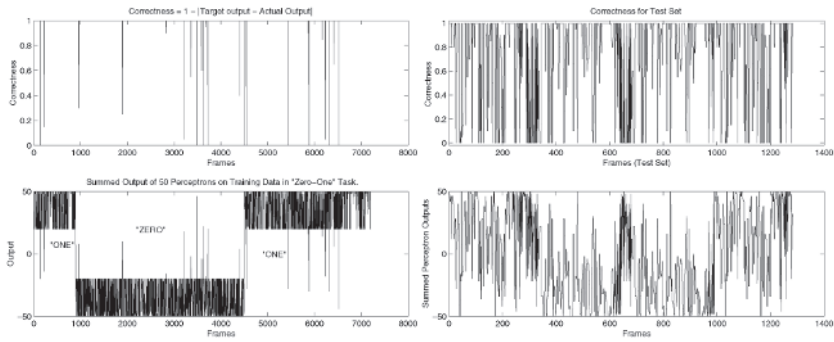


Fig. 8. Left: Perfect classification of training data. Right: Generalisation errors occur in apx 35% of frames.

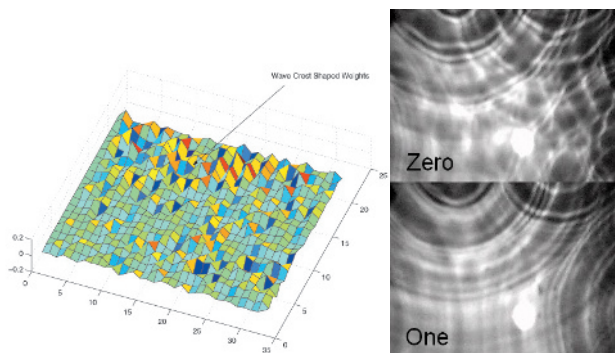


Fig. 9. Left: Mean weight matrix. Right: Typical frames for “Zero” and “One” wave patterns.

but where connectivity falls off with distance [11]. We hypothesise that water waves exhibit similar “functional connectivity”. The uncertainty about the state of subset X_j of the system X , which is accounted for by the state of the rest of the system ($X - X_j$) is the mutual information, given by

$$MI(X_j; X - X_j) = H(X_j) + H(X - X_j) - H(X) \quad (1)$$

Complexity is the area under the graph of MI for subunit sizes from $k = 1$ to $k = N/2$ where N is the size of X [11]. We measured the complexity of water under differing extents of “spontaneous activity”, i.e. due to random stimulation by 0 to 8 motors. Figure 10 shows that complexity increases in a staggered fashion with increasing numbers of motors. This is due to peculiarities in the way individual motors affect the water, some making large ripples and some making relatively small ripples. An interesting feature is that the system remains complex with increasing numbers of motors. Even with 8 motors stimulating the water the system does not become chaotic, that is, by observing less than 50 randomly chosen pixels of the image of water, we gain all the information that is stored in the entire array of pixels.

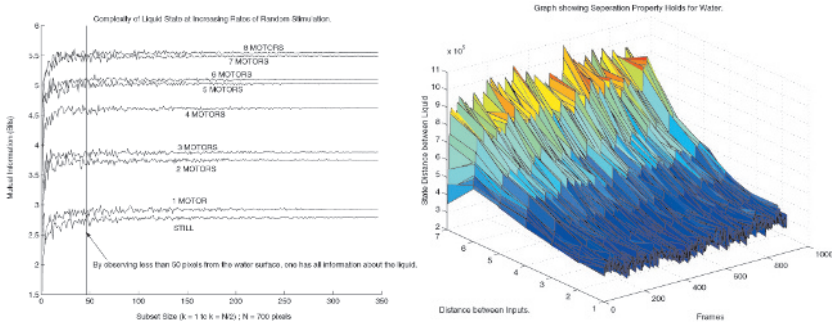


Fig. 10. Left: Complexity of water for increasing amounts of random stimulation. Right: Approximated separation property of water, where input distance is defined as the difference in number of motors stimulating water, and the state distance of the water is defined as the Euclidean distance of pixel values in the L-2 norm.

4.2 Separation Property

We approximated the separation property of water by the following method. We used the same data set as in part 4.1. The Euclidean distance between inputs was defined as the difference between numbers of motors stimulating the water between two trials. Figure 10 shows that the Euclidean distance between the state of the liquid increases linearly with Euclidean distance between inputs, just as in Maass’ LSM. We did not test the separation property for more subtle measures of input distances such as the patterned motor activity used in the speech discrimination task for example. The separation property is unlikely to hold for as wide a range of measures of input distance as in Maass’ LSM.

5 Conclusion

Our results show that the properties of a natural dynamical system (water) can be harnessed to solve nonlinear pattern recognition problems and that a set of simple linear readout elements will suffice to make the classification. This is possible without exhaustive tweaking of system parameters or careful efforts to decrease the level of noise in the system. Further work will use artificial evolution to create self-organising “neural networks” which are made of and exploit the epigenetic physical properties of biological materials [12].

Acknowledgements. Special thanks to Inman Harvey, Phil Husbands, Ezequiel Di Paolo, Emmet Spier, Bill Bigge, Aisha Thorn, Hanneke De Jaegher, Mike Beaton.

References

1. Maass, W., T. Natschlager, et al. (2002). “Real-Time Computing Without Stable States : A New Framework for Neural Computation Based on Perturbations.” *Neural Computation* 14: 2531–2560.
2. Hopfield, J. J. and C. D. Brody (2000). “What is a moment? ”Cortical” sensory integration over a brief interval.” *PNAS* 97(25): 13919–13924.
3. Auer, P., H. Burgsteiner, et al. (2002). “The p-Delta Learning Rule for Parallel Perceptrons.” submitted for publication.
<http://www.cis.tugraz.at/igi/pauer/publications.html>
4. Adamatzky, A. (2001). “Computing in nonlinear media: make waves, study collisions.” *Lecture Notes in Artificial Intelligence* 2159: 1–11.
5. Adamatzky, A., De Lacy Costello, B.P.J (2002). “Experimental logical gates in a reaction-difusion medium : The XOR gate and beyond.” *Physical Review E* 66 046112(2002).
6. Goldenholz, D. (2002). “Liquid Computing : A Real Effect.” BE707 Final Project. Boston, Boston University School of Medicine.
<http://www.lsm.tugraz.at/people.html>.
7. MacLennan, B. (1999). “Field Computation in Natural and Artificial Intelligence”. Knoxville, University of Tennessee. Report UT-CS-99-422.
<http://www.cs.utk.edu/mclennan/fieldcomp.html>
8. McCulloch, W. S. and W. Pitts (1943). “A logical calculus of the ideas immanent in nervous activity.” *Bulletin of Mathematical Biophysics* 5: 115–133.
9. Freeman, W. (2000). “Mesoscopic Neurodynamics : From neuron to brain.” *Journal of Physiology (Paris)* 94: 303–322.
10. Walmsley, I. (2001). “Computing with interference: All-optical single-query 50-element database search.” *Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science*, Baltimore, Maryland.
11. Tononi, E., Sporns (1998). “Complexity and Coherency: integrating information in the brain.” *Trends in Cognitive Sciences* 2(12): 474–483.
12. Muller, G. and S. Newman (2003). *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*, MIT Press.