Coding Mental States from EEG Signals and Evaluating their Integrated Information Content: a Computational Intelligence Approach

Rita Pizzi and Marialessia Musumeci

Abstract— The paper presents a method to identify and code mental states from EEG signals, performing their dynamical analysis by means of an Artificial Neural Network. The method has been tested on signals from a 14 electrodes EEG system connected to immersive glasses that allow a realistic audiovisual experience. A software procedure synchronizes the acquired signals with the sensory experiences presented in a video. A suitable Artificial Neural Network detects and codifies the chaotic attractors signals related to the sensory and cognitive events. The analysis shows that the binary codes corresponding to similar cognitive and perceptive stimuli are similar, and well differentiated from the codes corresponding to different stimuli. The dynamical attractors corresponding to each mental state are submitted to a procedure that evaluates their Integrated Information content in the qualia space.

Keywords— Artificial Neural Networks, EEG signals, Cognition, Chaotic Attractors, Integrated Information Theory, Qualia

I. INTRODUCTION

THIS Neuroscience and in particular the consciousness studies constitute one of the most fascinating and challenging issues of modern science.

Several authors have tried to characterize the general properties of consciousness: among many others, [1]-[5].

There is a broad consensus on the fundamental characteristics of consciousness: its subjective, qualitative nature, continuing over time, in the sense that memory connects the consciousness of the present with the consciousness of the past; its multisensory nature, connected to the processes of thought, emotion, memory, imagination, language and action planning.

Most of the neurobiological theories on consciousness show that cortical and thalamus activity plays a critical role and provides much of the content of consciousness. Most of these theories are also based on the assumption that the neural correlates of consciousness (NCCs) are a functional complex of neuronal cells, but it is not yet clear which neurons, cortical regions or connections they involve.

The search for NCCs is one of the most difficult challenges of modern neuroscience. Despite of many new theories, analytical methods and tools, the subjective experience still hides from a precise neurophysiological identification [6]-[8].

A recent prominent theory is the Integrated Information Theory (IIT) developed by G. Tononi and M. Edelman, [5],[9], [10]. It is based on the fundamental principle that consciousness stems from the rapid integration of a large amount of information into a dynamical nucleus of strongly interacting elements; interconnections between regions of the thalamo-cortical system mediate this rapid integration.

The authors refer to the mathematical theory of information proposed by C. Shannon and W. Weaver [11]. According to this framework, information is defined as the reduction of uncertainty among a number of possible outcomes x of a random variable X when one of them occurs. Thus an increase of uncertainty corresponds to higher information, and the information content of x , $I(x)$, will be a decreasing function of its probability. Shannon showed that this function is expressed by

$$I(x)=-log_2 P(x)$$

where $P(x)$ is the probability that x occurs.

Entropy of the random variable X is defined as the expected value of the information content of X (i.e. its average information content)

$$H(X)=E(I(X)) \,.$$

Thus Entropy can be defined as a measure of the uncertainty associated with X.

Given two subsets A and B defining a single bipartition of a system X, Mutual Information measures the uncertainty of A that is accounted for by the state of B, and is defined as

$$MI(A;\, B) = H(A) + H(B) - H(AB) \,.$$

The Effective Information of a system measures the extent to which its repertoire of possible states is differentiated in

R. Pizzi is Senior Researcher at the Dept. of Computer Science, Università degli Studi di Milano, via Bramante 65 , 26013 Crema CR, Italy (corresponding author: tel. +390250330072; e-mail: rita.pizzi@unimi.it; website: http://www.di.unimi.it/pizzi).

M. Musumeci is PhD student at the Department of Computer Science, Università degli Studi di Milano (e-mail: marialessia.musumeci@unimi.it).

response to all possible inputs. Effective Information is calculated as the Mutual Information across a partition when the outputs from one subset have maximum Entropy.

As before mentioned, the IIT hypothesizes that consciousness corresponds to the capacity of the system to integrate information, and this measure is indicated by Φ.

Φ is defined as the Effective Information across the weakest link of the system, i.e. the Minimum Information Bipartition. The Minimum Information Bipartition is the partition of the system for which the Effective Information is lowest.

To summarize, in order to calculate the integrated information of a system Φ (thus a measure of consciousness, as intended in the IIT framework) it is sufficient to calculate the integrated information between two partitions of the system, among all the possible ones, which have the lesser amount of effective information between them. A high value Φ will denote highly structured complexity.

In the brain, the thalamo-cortical system can be described as a single large highly complex system whereas, on the contrary, the cerebellum consists of a large number of very small complexes, each corresponding to a single module and thus having a very low complexity. In terms of complexity, the differences between brain and cerebellum are therefore not the amount of effective information related to the repository of possible states that characterize each system, but rather the level of integration of the information contained therein. Hence the flow of information between two parts of the same system must be considered.

But once defined a formal way to measure Φ and once identified a brain system that may represent a good candidate to generate integrated information, a method to represent complexity and integration in brain structures in such a way as to quantify Φ from real data must be developed.

One of the methods currently studied to analyze complexity in brain structures is to study the brain as a dynamical system.

Brain dynamics refers typically to the dynamics of neuronal populations, networks or columns within cortical areas. It is characterized by its high complexity, often involving oscillations at different frequencies and amplitudes, perhaps interrupted by chaotic or pseudo-chaotic irregular behaviour.

Synchronization among groups of neurons were first discovered in the olfactory system [12],[13], but has also been demonstrated in other brain structures, such as the hippocampus [14]-[18] and the visual cortex [19],[20], where the oscillations tend to synchronize in phase.

Synchronous oscillations can occur in nearby neurons, but also over considerable distances across spatially separate columns [20] and even between cortical areas [19],[21].

According to IIT, several aspects of the organization of the cortico-thalamic system and of transient attractor dynamics appear well suited to information integration.

It has been recognized that the massive interconnectivity within and among cortical areas (and with thalamus) provides an ideal substrate for cooperative dynamics among distributed neurons [22]. A plausible scenario for characterizing such dynamics is in terms of *transient attractors.*

In fact neurons in the cortico-thalamic system seem to behave in such a way as to ensure the rapid emergence of firing patterns that are distributed over wide regions of the cortex, where some neurons are strongly activated, and many more are deactivated. These firing patterns remain stable (hence they form attractors) over a time scale of tens/hundreds of milliseconds, but then rapidly dissolve (hence the attractors are transient), to make room for another transient attractor.
Attractors have been indicated in the form of binary strings (e.g. in a Hopfield network consisting of 8 elements with 6 embedded attractors, the attractors are indicated with 00001111, 00110011, 01010101, and their mirror images.)

Metastable systems, namely dynamical configurations that constitute non-fixed-point attractors, are good candidates to form a class of systems with high Φ [9],[10],[23].

Our approach stems from the wide literature mentioned above. By means of a novel self-organizing ANN, called ITSOM, we show how the dynamical analysis of neural signals may highlight the existence of chaotic attractors, differentiated depending on the cognitive states, that outlines the attractors in which the corresponding dynamical system is evolving.

This model has been used in the past in other researches of our group, allowing to analyze multiple neural signals and to identify complex patterns corresponding to specific dynamical attractors in signals [24],[25].

If the attractors show to be chaotic, this means that the neural signals are individually self-organized and, when analyzing more signals together, that there is a form of coherence between signals. . The ANN can also highlight the time course of this form of coherence and identify different attractors with a unique code. The ANN allows to attribute the same codes to similar but not identical brain events, reaching the necessary range of flexibility.

II. METHODS

The Self-Organizing Map (SOM) [26],[27] features are well known. The SOM is essentially a classifier that performs a vector quantization, that is a mapping from a space with many dimensions to a space with a smaller number of dimensions, preserving the initial topology.

It is constituted by an input layer (in this case the signal that flows in time in the layer, one sample for each neuron) and a competitive layer, where the neuron closest to the input "wins" and is modified in such a way that the new adjusted weight for the node is equal to the old weight, plus a fraction of the difference between the old weight and the input vector:

$$w_{inew} = w_{iold} + \alpha(x - w_{iold})z_i$$

where $0 < \alpha < 1$ slowly decreases over time with the law

$$\alpha(t) = \alpha[1 - t/\delta]$$

where δ is a suitable constant, being $z_i \neq 0$ only for the winning neuron.

Then the network cycles adapting itself up to a stable state.

As above mentioned, the ANN model adopted in this research, named ITSOM (Inductive Tracing Self-Organizing Map), is especially suited for identifying structures in temporal series.

The ITSOM architecture stems from the SOM architecture but is based on the observation that the time sequence of the SOM winning weights tends to repeat itself, constituting chaotic attractors that are isomorphic to the attractors of the signal time series, and characterize univocally the input signal that produces them.

The ITSOM network memorizes the time series of the winning nodes, and this sequence makes it possible to classify the corresponding input value much more finely than with a SOM.

A detailed description of the ITSOM's architecture is reported in [28] .

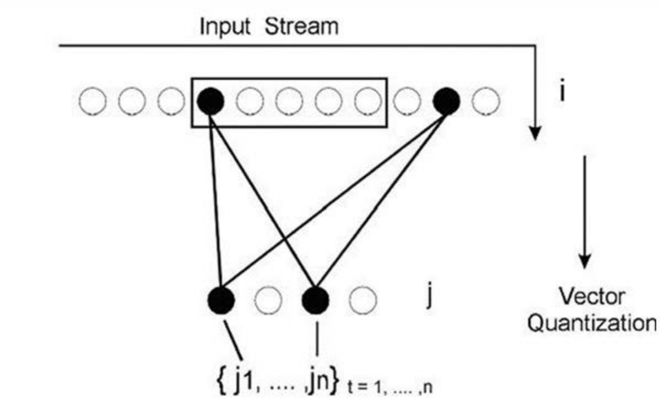


Fig. 1. the ITSOM structure: The sequence of the ANN winning nodes tends to repeat itself creating a cahotic time series that carachterizes the input signal.

A crucial feature of the ITSOM is that the cyclic configurations stabilize within a small number of epochs, that makes this model very effective for real-time applications.

The cumulative scores for each input are normalized according to the distribution of the standardized variable z given by

$$z = \frac{(x - \mu)}{\sigma}$$

where μ is the average of the scores on the neurons of the competitive layer and σ is the standard deviation.

Once set a threshold $0 < \tau \leq 1$, which therefore constitutes one of the parameters of this type of network, we put

$$z = 1 \quad for\ z > \tau$$
$$z = 0 \quad for\ z \leq \tau$$

In this way, each configuration of winning neurons is represented by a binary number formed by as many ones and zeros as many the output layer neurons.

Due to the existence of the threshold, the z-scores coincide when the series of winning sequences are approximately similar. Then the task of comparing z-scores becomes straightforward and allows us to identify similar or identical input patterns.

Analyzing the signals by means of the ITSOM network, it can be shown that attractors are labeled with a binary code that identifies them univocally, but the flexibility of the ANN allows to attribute the same codes to similar dynamical events: this is an important issue, as of course neural signals are never identical even when the stimulus that influences them is the same [29].

In this way we obtain a fine classification of the signal on the basis of its dynamical self-organization in time.

III. THE EXPERIMENTAL PHASE

In this study we processed signals from a 14 electrodes of the EMOTIV+ wireless EEG system [30] (Fig. 2), connected to immersive glasses that allow a realistic audiovisual experience.

Fig. 2. The Emotiv+ system and the 14 electrodes

The performances of the EMOTIV+ headset were evaluated in literature as equal to - or better than - a research EEG headsets [31].

The subject wears both glasses and EEG headset. A video administers sensory and cognitive stimuli, each one lasting 10 s, followed by a 5 s black stimulus, as a function of control and reset (Fig. 3). We chose different colors, colored images and written words repeating the colored stimuli.

A procedure developed in [32] synchronizes the acquired signals with the various sensory and cognitive experiences presented in the video.

At the end of the experiment, signals are recorded and the analysis procedure is applied.

We chose in particular to process four electrodes (T8, P7, O1, F7) (Fig. 4) as the most interesting in relationship with the chosen stimulations. In fact F7 is involved in cognitive control, T8 in episodic memory, P7 in visuospatial processing and the O1 main functional area is the primary visual cortex.

The frequency analyzed were Beta (between 12.5 and 30 Hz) and Gamma (>30 Hz).

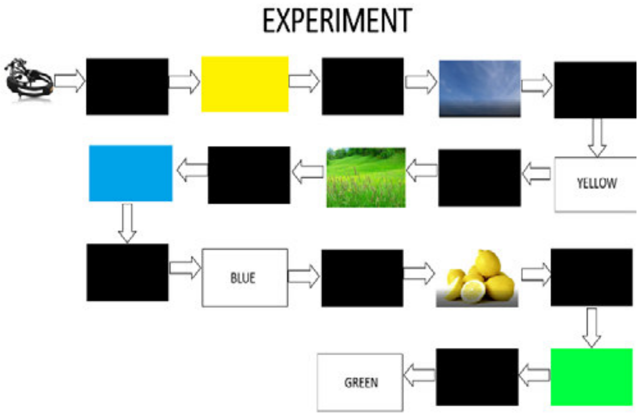


Fig. 3. The video administered to the subjects. The sensory and cognitive stimuli last 10 s and are followed by a black stimulus lasting 5 s.

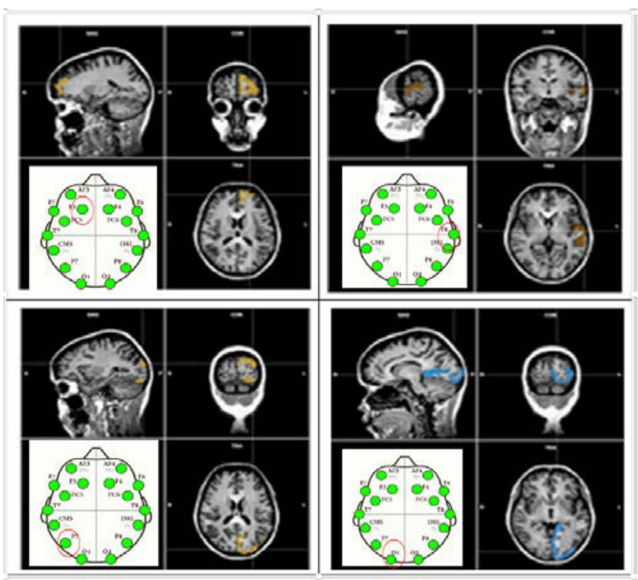


Fig. 4 . The 4 channels chosen for the analysis: T8, P7, O1, F7

Aim of the analysis is to test if similar stimuli give rise to chaotic attractors identified with identical or similar codes.

ITSOM can process both individual signals and many signals simultaneously, highlighting the attractors in which the corresponding dynamical system is evolving. If the attractors are chaotic, this means that the signals are individually self-organized or, if you examine more signals together, that there is a form of coherence between signals. The ANN can also highlight the time course of this form of coherence.

Once the time series of the attractors is available, it is also possible to quantify these complex dynamical events with many parameters useful to compare the dynamical events corresponding to different kinds of stimulations.

IV. RESULTS AND CONCLUSIONS

Signals were acquired from seven subjects: results are not comparable, as by definition each subjective experience is different from subject to subject. But the analysis of the binary codes resulted from the ITSOM processing shows the constant evidence that in any subject's signals most binary codes are identical or similar for similar patterns, and different for different patterns.

The figures show the analysis of the signals from one of the subjects. In particular, the shown analysis concerns the Gamma band of the T8 electrode. Gamma band gave the best results. This can be a further confirmation that the most prevalent physiological candidate for a key role in consciousness is synchronized neuronal activity in the gamma frequency band, approximately 35-45 Hz. This hypothesis is supported by a number of studies that have highlighted a widespread gamma synchronization in the magnetoencephalogram (MEG) in REM sleep and sleep state [33] and auditory evoked potential, used as marker of the state of consciousness in anesthesia studies [34]. Engel and Singer have suggested that synchronization can play a role in all the underlying processes of consciousness: arousal, sensory segmentation, selective focus, memory, and even in higher cognitive processes such as motivation, action planning, and symbolic processing [21].

In Fig. 5a, 5b, 5c the first columns show the sensory and cognitive stimuli, the second columns show the binary code resulted from the ANN processing, the third columns show the attractors generated by the dynamics of the sequence of ITSOM winning neurons: the figure represents a snapshot of movies that show a typical chaotic path.

To summarize the results, comparing the stimuli, the codes in Table 1 are obtained, clearly highlighting how similar stimuli give rise to similar codes, that result to be quite different from the codes obtained by different stimuli.

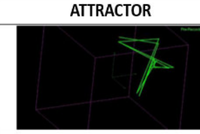
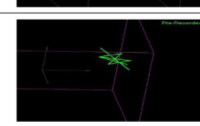
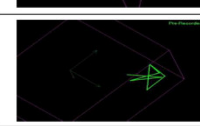
GAMMA BAND (>30 Hz) – *ELECTRODE T8 Temporal Lobe*		
STIMULI	CODE	ATTRACTOR
	1100101111	
	1100001111	
YELLOW	11000111111	

Fig. 5a. Binary codes and attractors of Yellow or similar stimuli

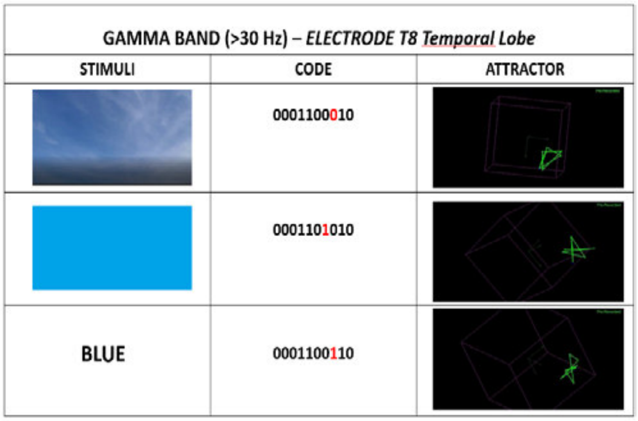


Fig. 5b. Binary codes and attractors of Blue or similar stimuli

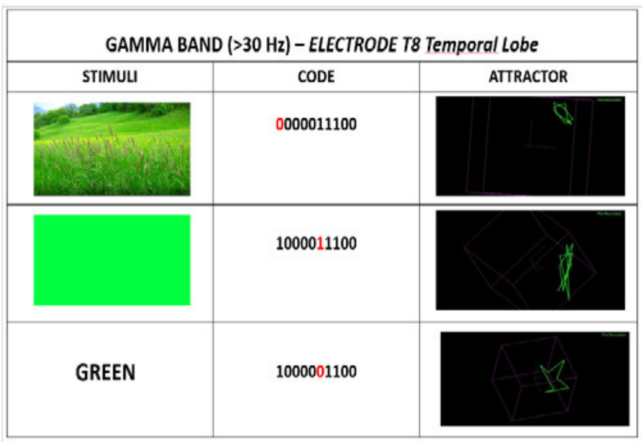


Fig. 5c. Binary codes and attractors of Green or similar stimuli

YELLOW-> 1100101111	BLUE-> 0001101010	GREEN-> 1000011100
LEMONS-> 1100001111	SKY-> 0001100010	MEADOW-> 0000011100
WRITTEN YELLOW-> 1100001111	WRITTEN BLUE-> 0001100110	WRITTEN GREEN-> 1000001100

Table 1. Summary of the results. Codes of similar stimuli are similar, codes of different stimuli are quite different.

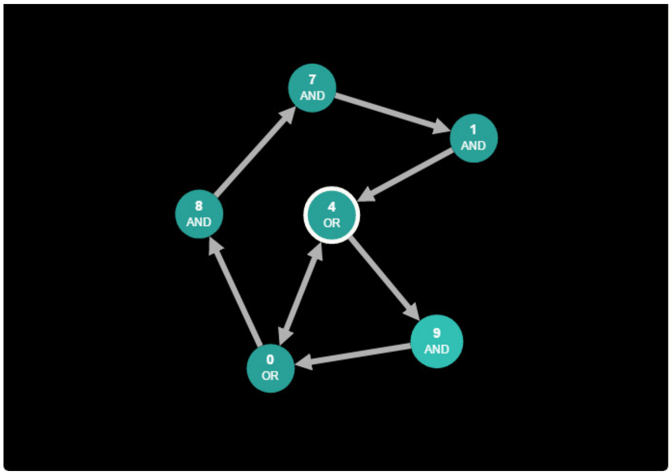
Using the tool available at the website of the Center of Sleep and Consciousness of the University of Wisconsin [35],[36], that allows to explore the Information Integration theoretical framework, we have been able to calculate the Φ value of the specific patterns through their related dynamical attractors.
The summary of results in sketched in Table 2.
Φ represents the integration at a system level, whereas any φ *measures the integration at the mechanism (subsystem) level.*
The dynamical representation in the concept (qualia) space of some of the patterns is reported in Fig. 6 and Fig. 7.

Pattern	Φ	Σφ
YELLOW	0.08323	1.6131
BLUE	0.08323	0.69643
GREEN	0.21528	0.91667
LEMON	0.21528	1.41667
SKY	0.08323	0.69643
MEADOW	0.21528	0.91667
WRITTEN YELLOW	0.21528	0.91667
WRITTEN BLUE	0.21528	1.41667
WRITTEN GREEN	0.21528	1.91667

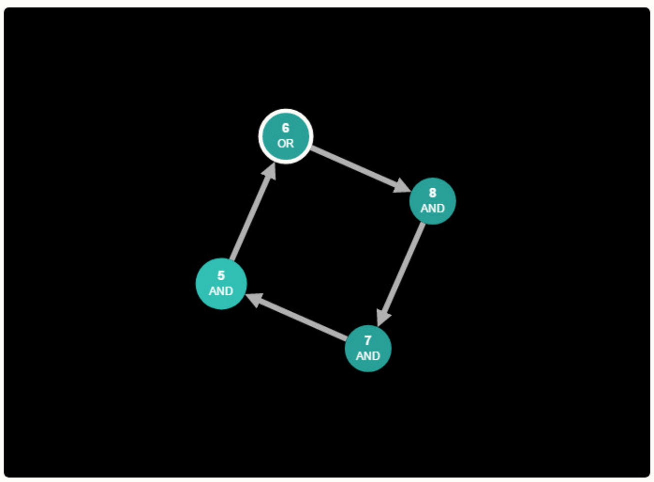
Table 2. Integrated Information Calculus

In conclusion, for all the color stimulations the Φ value was equal to 0.08323, except for the Green color that had a Φ value equal to 0.21528. The other stimulations had a Φ value higher than the pure colors and equal to 0.21528: in line with the IIT, this is correct as they have not only sensory but also cognitive contents, thus should involve more neural structures and can be considered more complex. Although the Φ value of stimulation patterns of information content with equivalent complexity coincide, their specific information contents are diverse and composed by subsystems with different values.

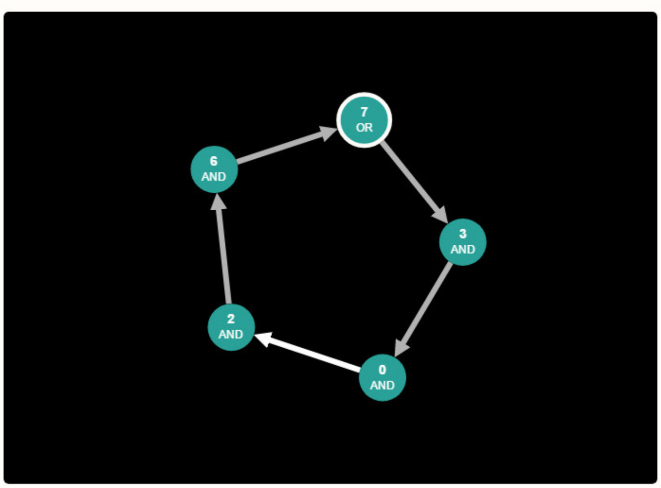
We would be tempted to state that these codes can be a way to identify qualia, i.e. the subjective and qualitative experience of mental conscious states and of their neural correlates [23], as there is an extremely high number of possible binary codes, but we can distinguish a set of dynamical states with unique codes that we may call “qualia codes”.



Attractor of : Yellow

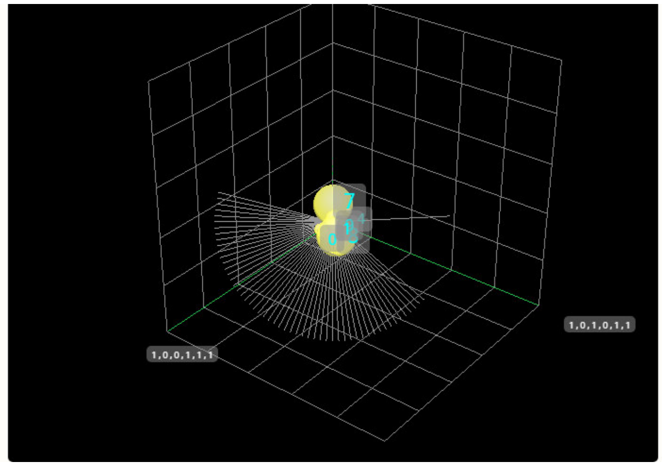


Attractor of: Lemon

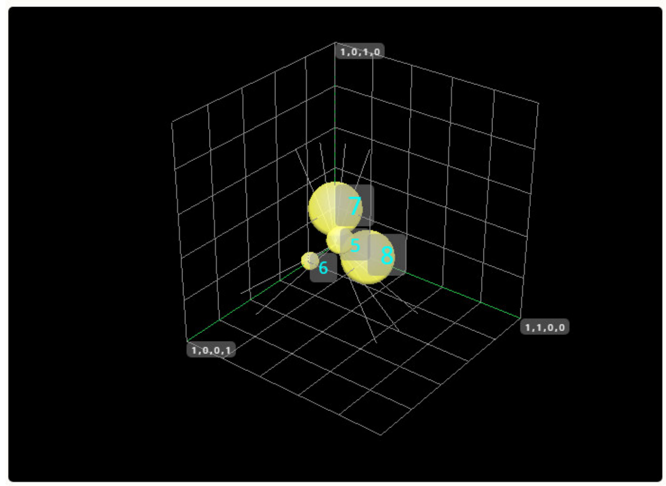


Attractor of: Written Green

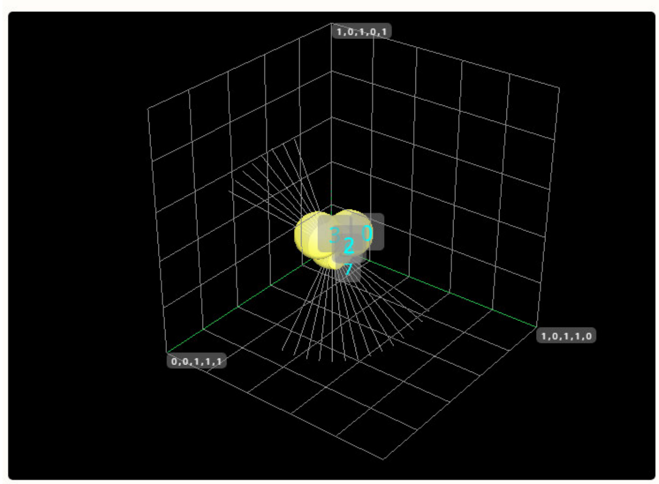
Fig. 6. Dynamical representations of some of the patterns



Conceptual structure of: Yellow



Conceptual structure of: Lemon



Conceptual structure of: Written Green

Fig. 7. Conceptual structures of the patterns in Fig. 5

In this case, as the size of the competitive layer is 10, the number of possible codes is $2^{10} = 1024$. For more complex patterns, the choice of a higher number of codes would be more suitable. Future developments of this research aim to identify more numerous and complex sensory and cognitive stimuli. Currently we are experimenting a new set of visual, auditory and cognitive stimuli, overlapping and comparing them with emotional stimuli, with promising results.

The IIT dynamical approach described in [9],[10],[23] does not fully specify yet the underlying dynamics of real signals and the way to identify it, due to the lack of a robust quantification method.

We hope that our contribution may be useful to go one step further towards the fine-grained discrimination of mental states by means of brain dynamics analysis, making it possible to quantify and evaluate their Integrated Information content.

References

[1] T. Nagel, "What is it like to be a bat?," in Nagel T. ed., *Mortal Questions*, Cambridge University Press, 1979, pp. 165–80.
[2] W. James, *The Principles of Psychology*, New York: Henry Holt, 1890.
[3] S. Greenfield, "How Might the Brain Generate Consciousness?," in Rose S. ed., *From Brain to Consciousness*, London: Allen Lane, pp. 210-270, 1998.
[4] J. Searle, "The Problem of Consciousness," *Social Research*, Springer, vol. 60, no. 1, 1993.
[5] G. Tononi, M. Edelman, "Consciousness and the integration of information in the brain," *Advances in Neurology*, vol. 77, no. 80, pp. 245, 1998.
[6] T. Metzinger, *Neural Correlates Consciousness: Empirical and Conceptual Question*, MIT Press, 2000.
[7] G. Tononi, C. Koch, "The Neural Correlates of Consciousness: an Update," *New York Academy of Science*, vol. 1124, pp. 239-261, 2008.
[8] C. Koch, G. M. Boly, "Neural Correlates of Conciousness: Progress and Problems," *Nature Reviews Neuroscience*, vol. 17, pp. 307-321, 2016.
[9] G. Tononi, D. Balduzzi, "Integrated Information in Discrete Dynamic Systems: Motivation and Theoretical Framework," *PLOS Comput. Biol.,* vol. 4, no. 100091, 2008.
[10] G. Tononi, "Consciousness as Integrated Information: a Provisional Manifesto," *Bio. Bull.*, vol. 215, no. 3, pp. 216-242, 2008.
[11] C. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379-423 & 623-656, 1948.
[12] E.D. Adrian, "Olfactory Reactions in the Brain of the Hedgehog," *The Journal Physiology*, vol. 100, no. 4, pp. 459-473, 1942.
[13] W. Freeman, *Neurodynamics: An Exploration in Mesoscopic Brain Dynamics*, Springer, 2000.
[14] J.D. Green, A. Arduini, "Hippocampal Electrical Activity in arousal," *J. Neurophys.*, vol. 17, n. 6, pp. 533-557, 1954.
[15] G. Buzsáki, Z. Horváth, R.Urioste, J .Hetke, K.Wise, "High-frequency Network Oscillation in the Hippocampus," *Science*, vol. 256, no. 5059, pp. 1025-7, 1992.
[16] M. Steriade, "Neurophysiologic Mechanisms of Non-Rapid Eye Movement (Resting) Sleep," in: Chokroverty S. ed., *Sleep Disorders Medicine*, Boston: Butterworth Heinemann, 1999.
[17] R. Llinás, "The Intrinsic Electrophysiological Properties of Mammalian Neurons: Insights into Central Nervous System Function," *Science*, vol. 242, no. 4886, pp. 1654-64, 1989.
[18] C. Pantev, T. Elbert, B. Lütkenhöner, *Oscillatory Event-Related Brain Dynamics*, New York: Springer, 1994.
[19] R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk, H. J. Reitboeck, "Coherent Oscillations: a Mechanism of Feature Linking in the Visual Cortex? Multiple Electrode and Correlation Analyses in the Cat," *Biol. Cybern.*, vol. 60, no. 2, pp. 121-130, 1998.
[20] C.M. Gray, W. Singer W., "Stimulus-specific Neuronal Oscillations in Orientation Columns of Cat Visual Cortex," *Proc. Natl. Acad. Sci. USA,* vol. 86, no. 5, pp. 1698-702, 1999.
[21] A.K. Engel, P. Fries, P.R. Roelfsema, P. Koing, W. Singer, "Temporal Binding, Binocular Rivalry, and Consciousness," *Conscious Cogn.,* vol. 8, no. 2, pp. 128-151, 1999.
[22] F. Crick, C. Koch, "A Framework for Consciousness," *Nat. Neurosci*. vol. 6, no. 2, pp. 119-126, 2003.
[23] G. Tononi, "Integrated Information Theory of Consciousness: an Updated Account," *Archives Italiennes de Biologie* , vol. 150, no. 4, pp. 293-329, 2012.
[24] R. Pizzi, G. Cino, F. Gelain, D. Rossetti, A. Vescovi, "Learning in Human Neural Networks on Microelectrode Arrays," *Biosystems*, vol. 88, no. 1-2, pp. 1-15, 2007.
[25] R. Pizzi, G. Cino, F. Gelain, D. Rossetti, A. Vescovi, W. Baer, "A Cultured Human Neural Network Operates a Robotic Actuator," *Biosystems*, vol. 95, no. 2, pp. 137-144, 2009.
[26] T. Kohonen, *Self-Organisation and Association Memory*, Springer Verlag, 1983.
[27] T. Kohonen, "Physiological Interpretation of the Self-Organizing Map Algorithm," *Neural Networks*, vol. 6, no. 7, pp. 895-905, 1993.
[28] R. Pizzi, M. Musumeci, *Artificial Neural Networks, Dynamical Systems and Self- Organization*, Createspace, 2017.
[29] R. Pizzi, M. De Curtis, C. Dickson, "Evidence of Chaotic Attractors in Cortical Fast Oscillations Tested by an Artificial Neural Network, " in J. Kacprzyk, ed., *Advances in Soft Computing*, Heidelberg: Physica Verlag, 2003, pp. 11-22,
[30] Emotiv Inc., San Francisco, CA [Online]. Available: https:/www.emotiv.com.
[31] N.A. Badcock,K. A. Preece, B. De Wit, K. Gleen, N. Fieder, J. Thie, G. McArthur, "Validation of the Emotiv EPOC EEG System for Research Quality Auditory Event-Related Potentials in Children," *PeerJ*, vol. 3, e907, 2015.
[32] *MATLAB and Statistics Toolbox Release 2012a*, MATLAB, The MathWorks, Natick, MA.
[33] R. Llinas, U. Ribary, "Coherent 40-Hz Oscillation Characterizes Dream State in Humans," *Proc. Natl. Acad. Sci. USA*, vol. 90, pp. 2078–81, 2003.
[34] R. Munglani, J. Andrade, D. J. Sapsford, A. Baddeley, J.G. Jones, "A Measure of Consciousness and Memory During Isoflurane Administration: the Coherent Frequency," *Br. J. Anaesth.*, vol. 71, no. 5, pp. 633–41, 1993.
[35] *Integrated Information Theory,* Center for Sleep and Consciousness at the University of Wisconsin–Madison [Online]. Available: http://integratedinformationtheory.org.
[36] M. Oizumi, L. Albantakis, G. Tononi , "From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0," *PLOS Comput. Biol.,* vol. 10, no. 5, e1003588, 2014.