Enhancing dynamical signatures of complex systems through symbolic computation

Alberto Porta$^{1,2}$, Mathias Baumert$^3$, Dirk Cysarz$^{4,5}$ and Niels Wessel$^6$

$^1$Department of Biomedical Sciences for Health, University of Milan, Milan, Italy
$^2$IRCCS Galeazzi Orthopedic Institute, Milan, Italy
$^3$School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, South Australia, Australia
$^4$Integrated Curriculum for Anthroposophic Medicine, and
$^5$Institute of Integrative Medicine, University of Witten/Herdecke, Witten, Germany
$^6$Department of Physics, Humboldt-Universität zu Berlin, Berlin, Germany

The most pressing challenges that modern signal processing is facing are (i) to provide a description of the behaviour of complex systems formed by interacting subunits without modelling the functioning of single constituents and their relations, (ii) to establish fully multivariate frameworks for quantifying dependences among signals along predefined temporal directions, (iii) to disentangle dynamical signatures pertaining to the functioning of the system constituents, while preserving the possibility to describe the activity of the whole organism (i.e. coordinated behaviour), (iv) to interpret non-stationarities as a sequence of quasi-stationary states reflecting the determinism of the underlying system, (v) to extract and classify patterns associated to system conditions and specific behaviours, (vi) to be robust against noise affecting traces recorded in real experimental settings, and (vii) to be computationally efficient to promote applications in real time.

Symbolic signal processing, defined as a set of model-free techniques transforming a set of contemporaneously recorded signals (i.e. a multivariate recording) into sequences of symbols with the specific aim of enhancing distinct features while discarding unessential details, can naturally deal with all the above-mentioned challenges. Indeed, it is a data-driven approach focusing
on feature extraction and pattern classification. The coarse graining schemes, usually used by symbolization procedures, render symbolic methods particularly efficient in limiting the effect of noise. Moreover, the techniques employed to group symbols into more complex structures, usually called ‘words’, favour the extraction of patterns that can be easily linked to system states. The originality of the symbolic approach, lying in the process of pattern creation and redundancy reduction, allows the generation of new categories promoting the identification of relevant behaviours from real data, the description of high-level functions given lower-level activities, the association between patterns and mechanisms (or functions) and the detection of pathological situations leading to alteration of the physiological pattern distribution. Simple strategies for building patterns from a single signal can be supplemented by more complex approaches creating schemes involving several signals, thus allowing the simultaneous representation of functions at different levels of integration. In addition, given that symbolic processing is mainly based on the computation of pattern rates, relevant invariants of the dynamics, such as Shannon and conditional entropy, can be calculated as a by-product of the pattern classification procedure, thus simplifying the quantification of indexes describing system complexity [1]. In the presence of a multivariate set of interacting signals, the implementation of more complex classification schemes considering the joint and conditional occurrence of patterns allows the calculation of indexes assessing the degree of association among signals, such as mutual information, and markers of directionality of the interactions and the strength of the casual relations [2], such as the amount of information carried by a signal that cannot be derived from the entire multivariate set and the mutual information conditioned by a subset of the considered signals. The compact view over the system dynamics resulting from pattern definition, classification and redundancy reduction makes the estimation of the above-listed dynamical quantities highly efficient, thus reducing computational time, saving processing resources and enabling real-time applications of complexity analysis.

Numerous applications of symbolic analysis can be found in the neuroscience and cardiovascular physiology literature. In neuroscience, symbolic computation was exploited to decompose spontaneous multi-channel electroencephalographic recordings into time segments characterized by quasi-stationary field map configurations, i.e. the so-called brain microstates [3], to study the recurrence of microstates when brain processing activity is triggered by a stimulus [4], to detect determinism in perictal intracranial electroencephalographic signals [5], to predict epileptic seizures [6], to identify abnormal resting-state magnetoencephalographic activity in early-stage Parkinson’s disease patients [7], to characterize event-related potentials [8], to evaluate the complexity of the electroencephalographic responses to transcranial magnetic stimulation during pharmacological and pathological loss of consciousness [9], and to assess causal relations between signals recorded from different brain regions (i.e. the so-called effective connectivity) under different levels of consciousness [10] and before the development of epileptic seizure [11]. In cardiovascular physiology, symbolic computation was used to track the gradual shift of the sympathovagal balance towards a sympathetic activation and vagal withdrawal during graded orthostatic challenge [12], to monitor the complexity of the cardiac control [13], to follow the modification of maternal baroreflex regulation during gestation [14], to study the closed loop interactions between heart period and systolic arterial pressure [15] and to estimate the reduction of the strength of the cardiopulmonary interactions during active standing [16]. The application of these methods triggered more clinically oriented studies that improved detection of pathological states and risk stratification in cardiology [17,18]. However, given the generality of the symbolic approach, studies exploiting symbolic computation can be found virtually in any field of science even though applications are more frequent in neuroscience and cardiovascular physiology literature [19]. This Theme Issue was proposed to stimulate new investigations especially in domains that have been little penetrated by applications of symbolic dynamics.

The issue will present techniques to symbolize time series, create symbolic patterns, collate patterns into a small number of families and calculate dynamical invariants from symbolic series. Both univariate and multivariate applications will be covered stressing the ability of the
approach to assess complexity and describe interactions among system constituents. The final aim is to provide practical and creative strategies for tackling the dynamical complexity in contexts where little is known about the physics of the system and the relations among subunits. The main spotlights of this Theme Issue will be the formalization of various symbolization techniques and their application over a single time series, the extension of these procedures in the presence of a set of contemporaneously recorded signals (i.e. joint and conditional symbolization), the identification of procedures for the construction of sequences of symbols (i.e. patterns), the proposal of redundancy reduction strategies for grouping patterns into a small number of classes according to predetermined criteria, the calculation of dynamical invariants describing interactions among system components such as directionality indexes, coupling strength and information transfer from symbolic traces, and the variety of applications ranging from neurosciences to cardiovascular physiology, from spatially extended networks to applied linguistics, from combustion engines to psychiatry. The issue was compiled to convince the reader that symbolic analysis is much more than a tool for providing a rough coarse-grained representation of reality. Symbolic analysis is actually an original way to look at a dynamical system by taking care of its constituents without losing sight of its global behaviour.

This Theme Issue comprises 11 contributions from experts with a long-standing experience in symbolization and symbolic computation in different contexts. The contributions can be roughly divided into theoretically oriented studies, where the methodology is the focus and archetypal examples of applications are provided for illustrations, and problem-driven investigations, where the emphasis is on the application, whereas mathematical details are discussed without a systematic approach. It is worth stressing that the majority of contributions provide a comprehensive extension to the multivariate case [20–27], thus providing tools fully helpful in contexts where multiple traces are acquired.

Among the technically oriented contributions, Amigò et al. [20] review the methods based on ordinal symbolic statistics and apply them to electroencephalographic recordings during epileptic seizure; Lehnertz & Dickten [21] stress the possibility offered by symbolic computation to assess coupling strength and directionality given a multivariate set of signals and apply the approach to identify interdependences between brain regions forming an epileptic network; and beim Graben & Hutt [22] refine a previously proposed symbolic analysis approach for segmentation of event-related potentials into quasi-stationary states and provide an application to human language processing.

Among the application-driven contributions, Cysarz et al. [28] apply several symbolization strategies to spontaneous beat-to-beat changes of fetal heart period and test their performances in following the development of the fetus during gestation; Valencia et al. [29] show that symbolic indexes derived from the beat-to-beat series of heart period and ventricular repolarization duration separate ischaemic-dilated cardiomyopathy patients and healthy age-matched controls better than more traditional linear time and frequency domain markers; Schlemmer et al. [30] propose a method based on symbolization to study sleep stage transitions and their modifications in relation to ageing and sleep disorders; Daw et al. [23] provide an industry-driven application of symbolic analysis to control dilute combustion instabilities in a multi-cylinder spark-ignited engine; Baumert et al. [24] review applications exploiting joint symbolic analysis to identify the pathophysiological changes of specific cardiovascular control mechanisms such as cardiac baroreflex and cardiopulmonary reflexes; Schulz et al. [25] present an application of a high-resolution joint symbolic analysis to separate patients suffering from paranoid schizophrenia from their healthy first-degree relatives and age-matched controls; Porta et al. [26] propose a conditional joint symbolic technique to assess the nonlinear interferences of respiration over cardiac and sympathetic ‘baroreflexes’ and apply it to an experimental protocol designed to evoke postural syncope; and Lee et al. [27] review the basic principles of the neurobiology of consciousness, the problem of ‘covert consciousness’ (i.e. the presence of conscious experience coupled with the absence of responsiveness) and the recent advances in assessing levels of consciousness and cortical connectivity using symbolic dynamics and symbolic transfer entropy.
Given the possibility to extract features attributable to specific system constituents and to describe their interdependencies, symbolic analysis might allow an efficient identification of anomalous states owing to the impairment of a connection among subsystems (i.e. missed connectivity) and/or owing to the decline in performance of the subunits in any aggregates of components, ranging from genetic sequences to physiological systems, from neural assemblies to social groups, from cortical areas to chemical reactions, from groups of sentences to mechanical components. In the meanwhile, the possibility to describe the global functioning of the system guarantees the detection of abnormally coordinated or uncoordinated behaviours. The diverse, contemporaneous views at different levels of integration can empower diagnostic ability and improve prediction of outcomes in any field of science. In addition, given the high computational efficiency of symbolic methods, metrics for the assessment of dynamical system complexity and information flow among components grounded on symbolization techniques might be exploited to monitor the behaviour of the system and its constituents in real time. Symbolic indexes might provide a reliable alternative to markers of complexity and causality that can be derived from model-based multiple linear regression approaches in time and frequency domains [31–34]. More specifically in the medical field, we expect that symbolic computation will have a powerful impact in clinics in the coming years by playing a significant role in tailoring individual treatments, improving diagnostics and therapy, managing patient data and reducing the cost of healthcare systems via a more precise risk stratification.

References


