Multiregime dynamics: modeling and statistical tools

Abstract: In this paper, we introduce the formalism and terminology of multiregime dynamics for both theoretical modeling and data analysis. Coding is proposed as the appropriate tool for the analysis of such special type of dynamics, focusing upon switches between suitably defined dynamical regimes. Individually taken these switches often represent abrupt alterations in the qualitative features of the dynamic process. At times, however, they seem to be stringed together to show emerging (near) regularities and fluctuations. This opens new vistas upon applications to the analysis of the vector time series of socioeconomic models. In empirical applications, coding involves transformation of data into a sequence of symbols that is then analyzed with information-theoretic tools, so as to extract information about generating processes.

Keywords: Dynamic regimes; coded dynamics.
JEL Classification: C60, C61, E30, E32.

Juan-Gabriel Brida[1]
Lionello F. Punzo[2]

[1] Free University of Bozen-Bolzano; Via Sernesi 1, 39100 Bolzano (Italy); e-mail: JuanGabriel.Brida@unibz.it

[2] Department of Economics, University of Siena, Piazza S. Francesco 7, 53100 Siena (Italy); e-mail: punzo@unisi.it

Introduction

The original motivation for the modeling approach presented hereafter arose in one of the discoveries in recent economic literature, a motivation which is now also emerging in other social sciences: even very simple dynamic models can display a surprising variety of behaviors. To capture this common feature, it has become standard to allow for e.g. multiple and/or indeterminate equilibria (dense sets), endogenous cycles and irregular fluctuations, and the like. In many instances, this modeling innovation can really be imputed to an effort to introduce formally a notion (that of *regime*), which already belongs to the jargon of the more empirically oriented scientists. In fact, nearly always dynamics is assumed to be governed by a single, universal system of differential or difference equations, as this is a useful simplification. However, behavior may be governed by different dynamic laws depending upon the values of the state variables, and therefore their uniqueness cannot be assumed but explicitly proved. Then changes in regime are qualitative changes in the dynamics brought about by changes in the underlying rules. Such changes can be single events, of the type associated with the idea of structural changes, or they can be coordinated in a temporal sequence, following more or less regular patterns. This general occurrence is what we call *multiregime dynamics* (MRD). MRD reproduces the complex dynamics of the socioeconomic models, solving the unsolved problem of marrying structural change with adjustment dynamics and can be conceived as a dynamical system over a set of models, and such set as kind of menu of qualitative behaviors available to a dynamical process. Formalizing such dynamics requires introducing the notion of a state space over models, which is naturally discrete. Now we can focus upon regime switches as discontinuous changes or discrete jumps. It is to isolate this kind of dynamics that regimes are assigned distinct symbols from an adequate alphabet, and our description goes symbolic defining a coded dynamics. The use of such technique is the fundamental difference of our approach from the conventional one where state variables are real numbers.

Compared with the original dynamics, coded dynamics deliberately overlooks certain regularities amenable to treatment by the usual methods, to focus upon qualitative features that, given the short compass of our data, will easily tend to be attributed to random components while on the contrary they may exhibit recurring (hence, in principle endogenously driven) patterns. As the coding procedure is mechanical and does not imply any priory assumption about the underlying processes, its by-product is an indirect test of which model might be appropriate to a given data set. Coding dynamics appears to be the appropriate tool to analyze models with multiple regimes, and it can easily be adapted to a space where states are discrete to begin with, as it happens in analyzing the dynamics of alternative social outcomes as in repeated games, and other theoretical and empirical social research.

We show some of the uses of coded dynamics to represent regime dynamics for socioeconomic models admitting multiple regimes, and discuss its relation with mathematical symbolic dynamics. In empirical applications as we see them, the coding techniques of the type we employ, information--theoretic statistical tools and computational experiments play a key role: they provide heuristics tools and are tools to explore data information. The difficulty of clearly grounding the explanation of the large body of empirical evidence in established theoretical approaches, has introduced further items in the theoretical agenda. Thus, for instance, multi-regime dynamics can be appreciated as a way to tackle the difficult issue of separating deterministic from disturbance elements in typical economic time series. If we do not resort to ad hoc assumptions on the generating mechanism, irregularity (often associated with non-stationarity) is in principle expected to be the general rule rather than a temporary, casual deviation from some fixed equilibrium value. Generally it is very difficult to identify rules constraining the set of variables representing a dynamical process to simple reciprocal relations, which would reduce the mathematical size of the problem to a smaller space of variables as is in the cointegration approach. In MRD, irregularity, i.e. lack of well defined attractors and similar features are taken to be as typical, inherent features of an economy's evolution. They are attributed, fundamentally, to the coexistence of various inseparable layers of dynamics as well as to the dynamic interdependence between the various components of an economy. Thus, in comparing systems over time, and in analyzing a system's history, the issue becomes how to measure or index irregularity.

The paper is organized as follows. Section 2 introduces the notions of dynamic regime and MRD. There, we review and compare different strategies to construct partitions into regimes and motivate our own choice and implied representation of regime dynamics via coded dynamics. In section 3, we formalize the notions introduced in section 2. In section 4 we present two pedagogical or teach-yourself models to illustrate the use of coded dynamics and its overlapping with the mathematical proof technique called symbolic dynamics. We also introduce entropy as a measure of irregularity or complexity in regime dynamics. In section 5 statistical methods arising in a natural way in the context of data analysis and processing of multiregime phenomena, are reviewed. These methods (collectively spanning so called Symbolic Time Series Analysis) are the statistical counterpart of coded dynamics. Finally, section 6 shows how such tools can be used to address the inverse problem of reconstructing a multiregime dynamical model from measured time series data.

Regimes and multiregime dynamics

The term regime has a relatively long history in economics (and probably in other social sciences as well). It has been used implicitly or explicitly extensively in a variety of fields, with reference not only to methodological aspects but also to analytical and economic policy, and even political issues.[1] Still, the term is generally not uniquely or well defined and it stands for different things to the various authors who have been using it in different contexts. In macroeconomic modelling and related econometric literature it is however commonly employed to identify an attractor, in other words an isolated equilibrium state (often a trend) towards which the system shows some local when not global stability. (see Durlauf and Quah (1999))

We introduce *our* notion of regime first in an intuitive way. It is easier to define it indirectly, by defining a regime switch. This refers to an event whereby a qualitative change in the functional forms of a given model has to be considered, as a result of the examination of the behaviour of the represented system. In fact, system behavior predicted by any posited model always implies the definition of a regime, and sometimes more than one, this being the set of rules governing it depending upon where it is travelling in its state space. Calling on economic intuition, an economic regime is therefore a set of rules and/or institutions, which are assumed to govern the economy at a given time and therefore to account for the qualitative features of its observed or observable behaviors. Regime, therefore, is often taken to stand for such a set of rules/institutions; sometimes, however, it is taken to indicate certain resulting qualitative behaviors, and often it is confusedly used to indicate either or both of them. The implied hypothesis is that model and predicted dynamics are uniquely corresponding to one another, which however cannot be taken for granted, to that it is better as we do, to keep the two notions distinct, and remain firmly with the notion of regime as a model with its own state space, capable thus of displaying a whole set of behaviours some which may be exhibited also by other regimes, though over other domains.

The interesting case is, of course, when there are two or more of such regimes, which may be as mutually excluding alternatives, or else as potential components of a dynamical menu of non-exclusive alternatives. The issue, then, becomes how such regimes get stringed together into a path or trajectory. The history of a given system as a specific realization of the set of alternatives in the menu may now display qualitative change. In the conventional approach the possibility of a system to select a specific regime is not excluded, of course; beyond the identification of regime with attractor, which is not here, what is excluded is the possibility that our approach permits, of visiting different regimes at different times, and for different lengths of time. The system may go from one to another through structural change in a variety of fashions.

Changes in regime are basically of the nature of discontinuous jumps or switches in the space of models spanned by the

[1] These different notions of regime are critically reviewed in Bimonte et als. (2001).

dynamical menu. Often, they can be modeled as the result of the system reaching and overshooting certain critical or threshold values in the key state values and/or in its parameters. Thus, to define a regime we need to identify its domain, or the state space or set over which a given set of mathematical rules apply, and in particular its boarder or frontiers where those rules or local model ceases to be the governing rules. In other words, when a system admits the possibility of occupying alternative regimes, its state space has to be partitioned, to identify the distinct regimes' domains. Hence, next to the problem of model selection, the criterion for partition becomes crucial. In general, we may distinguish between *endogenous* and *exogenous* partitions, depending upon whether they are induced by some pre-selected property (and therefore vary with the dynamic process under scrutiny) or are theory-induced (hence, they are pre-empirical or predetermined). An endogenous partition is one that can be constructed from the equations of the model, without attaching any particular interpretation to the necessary partition. This is chosen on the basis of mathematical convenience and it either exists or not. For example, for one dimensional systems usually we distinguish regimes by the monotone branch of the map that represents the model. Endogenous partitions are often used by mathematicians as a tool to obtain proofs for properties of a dynamical system. As we show latter in this paper, the representative example of an endogenous partition is a Markov partition. In principle, exogenous partitions descend from an at least intuitive understanding of the underlying dynamics that can be conceived as a sequence of phases or regimes. These partitions are generally induced by economic reasoning alone and may or may not be convenient from the mathematical point of view. For any particular process, one introduces a taxonomy of regimes for the problem at hand, this being determined by the theoretical beliefs or by the adoption of one or more theoretical frameworks. Then, exogenous partitions do not depend upon any hypotheses as to the data generating (deterministic and/or stochastic) model. Examples of exogenous partitions are those defined by a threshold value given some economic interpretation. For example, economists often talk of high and low inflation and then, we can define a regime of high inflation if the variable i is bigger than a particular threshold i_0 (for example, 10% of inflation) and regime of low inflation if $i \leq i_0$. The multiregime environment used in Brida, Puchet and Punzo (2003) makes use of the two dimensional Framework Space (introduced in Boehm and Punzo (1992)) which is constructed as the *domain space of growth theories*. Its partition into six regime sub-domains corresponds to different classes of existing economic theories and it is therefore defined exogenously.

Hereafter, we focus upon the possibility of a system displaying overtime dynamic behaviors that are qualitatively different. One image taken from the phase portrait of one such system is therefore a path taking across trajectories in principle belonging to different regimes. Therefore, each path can be seen as constructed piecewise, by stringing together pieces of trajectories predicted by different models. Each model would be a kind of local representation of an overall dynamics, thus accounting for only part of the history of the process. Put differently, an actual or simulated history is interpreted as one specific realization of a collection of already available regimes, a time sequence of part-trajectories through regimes with their own timing and duration in some agreed clock. Any change of regime naturally signals some form of structural change, the sequential ordering of visited regimes and other parameters of the time dimension giving information relevant to understand which form.

Defining regimes and multiregime dynamics

In this section, we translate the intuitive notion of regime introduced earlier into a formal definition which identifies each regime with its *own* specific and technically *local* mathematical model. Such definition leads to the idea that a regime change takes place whenever a system moves in its state space from one region to another where a qualitatively distinct (in terms of functional form and/or key parameter values) dynamic law applies. Structural change is the discontinuity in dynamics observed while crossing the border between them, when the whole model describing system behaviour abruptly is altered. This we will see can be a singular non reversible event, or else be coordinated and reversible.

Definition *If D is a subset of* R^n, $\{D_1, D_2, \dots D_n\}$ *a partition of D and* $f_i : D_i \rightarrow D$ (i = 1, 2, …, n) *is a family of functions, then each pair* $R_i = (D_i, f_i)$ (i = 1, 2, …, n) *is a regime. Conversely, given a dynamical system (D, f) and a partition* $\{D_1, D_2, \dots, D_n\}$ *of the domain D, a regime is a pai* (D_i, f_i) (i = 1, 2, …, n) *where* $f_i = f|_{Di}$ *is the restriction of the function f to the se* D_i.[2]

This definition reflects the fact that different regimes are described as regions of a state space in which the state variables exhibit characteristic behaviors. The global dynamics of a multiregime model is therefore represented by a dynamical system (D, f), while dynamics within region D_i is represented by the difference equation $x_{t+1} = f_i(x_t)$ $x_t \in D_i$, f_i being the restriction of f to the domain D_i. Of course, when $n = 1$, we are in a standard one-regime situation handled implicitly by the majority of models. However, an *interesting* partition slices the state space into (at least) two non empty sets D_i and it has as many distinct restrictions. As $f_i(D_i)$ is not necessarily a subset of D_i, paths need not be confined or get trapped into the slice of the state space where they originate; on the contrary, they generally may traverse from one regime to another. The simultaneous presence of multiple regimes provides alternatives: they are available to so to say pick and choose to assemble or construct a system´s history. One such history can moreover be quite rich, as it results from the cross product of a twofold dynamics, one *within* a given regime and one *across* regimes, running at the same time though with different clocks. The former represents the behavior within any specific regime, the latter captures and formalizes the concept of regime switch. Their mixing can produce any kind of dynamic behavior, and in particular may at least partially account for non-stochastically determined irregularity. In this setting, conventional dynamics corresponds to the former and MRD focuses upon dynamics across regimes to study structural changes, how they may be concatenated, and to analyze various degrees of irregularity as they emerge from actual histories. MRD is, in other words, an explanatory non stochastic framework for large irregular and fluctuating behaviours of socioeconomic systems. Structural changes can be dealt with one at a time with conventional nonlinear or linear-stochastic methods. MRD seeks an *endogenous* explanation for their tendency to repeat themselves and for their often irregular concatenation. The key to our definition is that the partition of the state space into a finite number of regime domains goes together with multiple dynamic models in a specific way. We can think of each of the regimes $R_i = (D_i, f_i)$ $i = 1, 2, \dots, n$ as the mathematical representation of a whole class of models rather than of an individual one.

Dynamics across regimes is defined on a finite set of regimes (hence it has a discrete domain) while dynamics within regime i is defined on the state space domain D_i as usual. That MRD be defined over a space of dynamical models (instead of space of real-valued coordinates), goes beyond other similar definitions already available in the literature[3]. States which are of this type only need a finite number of symbols to be represented. To represent such MRD it is natural to resort to the idea of symbolically coding them. The coding procedure translates a classical trajectory in a given state space into a trajectory in the space of regimes, whatever the dimension k of the original space. Hence, some of the limits that are encountered in the analysis of multidimensional systems can in principle be overcome. This representation of a system's dynamics is obtained by associating a symbol with a chosen regime and then coding observed dynamics into a string of symbols. The arithmetical precision of conventional analysis gets lost. On the other hand, coding dynamics may be a way to disentangle complicated, irregular patterns, identify discontinuous jumps that represent structural changes from the smoother displacements within a given regime and, possibly, to check some form of regularity is hiding behind them.

Formally, if we define the index of regimes by the function

$\pi : D \rightarrow A = \{1, 2, \dots, n\}$ such that $\pi(x) = i$ if and only if $x \in D_i$ ($i = 1, \dots, n$)

then, the symbolic sequence

$$(s_n(x))_{n\in N} = \pi(x)\pi(f(x))\pi(f^2(x))\pi(f^3(x))\cdots = (\pi[f^n(x)])_{n\in N}$$

[2] As usual in Symbolic Dynamics theory - see e.g. Alligood et als. (1997) p. 125-126- in this paper we will abuse terminology in calling a partition of the state space D, a collection of subsets of D, which have pairwise disjoint interiors and whose union is D. For a well defined dynamic rule, at a point x such that $x \in D_i \cap D_j$, $f_i(x) = f_j(x)$ must hold, of course.

[3] Although similar in spirit, our multi-regime and Day's multi-phase dynamics (see Day (1994, 2000)) have different origins and motivations, this accounting also for their distinct evolution.

gives the representation of the system as a sequence of regimes. This representation is called *coded dynamics* and is related (and partly overlaps) with the mathematical branch called symbolic dynamics.[4] It implies a re-coding of paths of conventional dynamics, which are described in real number coordinates as the infinite sequence $x, f(x), f^2(x),\ldots, f^n(x),\ldots$ through an initial state x, into a symbolic string $s_0s_1s_2,\ldots s_n,\ldots$ of symbols from a chosen alphabet $A = \{1,2,\ldots,n)$. If we start with $f(x)$ then the symbolic sequence will be $\pi\,(f(x))\,\pi\,(f^2(x))\,\pi\,(f^3(x))\ldots$. So, the dynamics on the space of symbol sequences just consists of shifting to the right: $s_0s_1s_2\ldots s_n,\ldots \mapsto s_1s_2\ldots s_n,\ldots$. Then, with n regimes, regime dynamics is defined in a subset of the set Σ_n of all symbol sequences and the dynamics is given by the *shift map* $\sigma : \Sigma_n \rightarrow \Sigma_n$ that shifts all coordinates one side to the right. We often think of an element $s_0s_1s_2,\ldots s_n,\ldots$ of the full shift as a time series, with s_0 representing the actual regime and s_1, s_2, …, s_n, … its future. The action of the shift map is like a tick of the clock, moving us one step into the future. The shift map is trivial so all the structure of the regime dynamics is given by the space of symbolic sequences generated by the map f. The complexity of the regime dynamics of a particular model is measured by the "size" of the space of feasible symbolic trajectories. We can use the tools of Information Theory to study the set of feasible symbolic sequences and in some cases we can write out an *abstract automata* which generates the symbol sequences produced the regime dynamics. If the original partition is well chosen, every point in the map's state space has a unique string of symbols, and vice versa, and shifting a string to the right is equivalent to iterating the map. When this is the situation one can find *Markov partitions* where there is an almost one-to-one correspondence between continuous states and the symbol sequences they generate. In these important cases, we are in the field of *symbolic dynamics* and studying the symbolic sequences is completely equivalent to studying the original (pointwise) dynamics. Markov partitions are not easily obtained in practice and unfortunately, it is well known that it is possible, that the state space can be so partitioned, only for a comparatively small class of special cases. On the other hand, there are no maps where it is known that it is definitely impossible to construct a generating partition. But the coding procedure of transforming the trajectory $x, f(x), f^2(x)\ \ldots, f^n(x), \ldots$ into a symbolic sequence can also be done when the partition is not Markov. In this case, we say that we have a coded dynamics. The proximity between coded and symbolic dynamics often permits the use of well-established symbolic dynamics techniques, as we will illustrate in the next section. This proximity shows up in particular when a mathematically more appropriate partition can be used that is finer than the regime partition induced by economic reasoning alone. To generate the former, whenever possible, we have to resort to a cross product of economic and the mathematical criteria. In fact, while our chosen definition of regime implies a partition of the system's state space, the latter may be introduced without paying any attention to its economic significance, and the two need not be consistent to one another. Still, a regime classification on the basis of some specific economic motivation can be a reasonable starting point to try to construct a mathematically useful construction. Broadly speaking, symbolic dynamic techniques can be effectively used whenever economic-theory induced partitions are included in those demanded by the former. This happens whenever they satisfy the requirement of being e.g. a generating partition. That there are cases where this is possible, is a good reason for further investigation.

In the next section, we illustrate our approach, with examples that have only a pedagogical meaning. In these examples we can give a detailed representation of regime dynamics and we can compute the complexity of the models, measured by its entropy.

Some pedagogical examples

The set of dynamic equations describing implicitly the law f_i governing over an associated domain D_i in the system state space D, are normally derived as the reduced form of a structural model, inclusive of definitions and other fundamental relations. Often such reduced forms have provided one-dimensional dynamical systems. In the literature on complex and chaotic economics, which has blossomed in the 80s early 90s,

[4] In Brida, Puchet and Punzo (2003) there is an intuitive introduction to the formalism and terminology of coded dynamics in multi-regime models. For an introduction to symbolic dynamics, see Alligood et als. (1997).

versions of the unimodal map has been popular, as giving in the discrete formulation a wealth of complicated behaviors with very minimal mathematical troubles, but often little economic justification, though, so that they remained theoretical exercises and counterexamples to well- established results.[5] In this vein, though, they may make good pedagogical devices, given the simplicity of their setting vis-à-vis the complicated outcomes they may yield. However, the earliest historical example to illustrate our notion of regime dynamics comes from the nonlinear cycle theory. In fact, in such a theory that goes back to the codification by Schumpeter, an economic cycle is broken down into *phases* separated by down- and up-turns, or switching points representing the sudden inversion of expansion or contraction. Therefore, one can conceive of the classical approach to the explanation of cycles, (in particular of business cycles, as defined by Schumpeter), as the ancestors of MRD. The difference being substantial though: not only because cycle theory was essentially confined to one-dimensional dynamics, while MRD is not, but more fundamentally because the regimes were taken to be concatenated in a regular sequence, and for this reason they were aptly denominated *phases of the cycle*[6]. The class of unimodal one-dimensional systems in discrete time is quite large, of course. In this context its members are useful whenever they permit the identification of at least two, if not more, non overlapping domains into which to split their one-dimensional state space, an open interval in R, with their associated local laws. We introduce the tent map as one representative of such class of unimodal maps, in order to illustrate in this setting the type of exercise of going from the standard representation in the real number system to the symbolic representation, the advantages of doing so when the outcome can be very complex and varied. this is well illustrated indeed by an unrestricted tent map.

Let D be normalized to $I = [0, 1]$ and let $f : I \to I$ be the function defined by:

$$f(x) = \begin{cases} 2x \text{ for } 0 \le x \le \frac{1}{2} \\ 2 - 2x \text{ for } \frac{1}{2} \le x \le 1 \end{cases}$$

and we consider the dynamical system (f, [0, 1]). This function is usually referred as the tent map and is increasing in $[0,\frac{1}{2}]$ and decreasing in $[\frac{1}{2}, 1]$. Partitioning one dimensional maps at the critical points generally work fine. This give us a mathematical criterion for the division into regimes: one regime is $I_0 = [0, \frac{1}{2}, 1]$ and the other $I_1 = [\frac{1}{2}, 1]$. These regimes will be represented by symbols 0 and 1 respectively. The sets I_0 and I_1 have disjoint interiors and their intersection is the point $x = \frac{1}{2}$. One such partition into domains, which is determined functionally on the basis of properties of the map, is, of course, an example of endogenous partition as defined above; it is also a Markov partition as it produces a conjugacy from the shift map σ on Σ_2 to f on I. Through such partition, and the properties of the tent map itself, in fact, one can show that a new (symbolic) representation of the tent map dynamical system by the shift map is obtained where all fundamental topological properties of the original system are preserved. Fig. 1 illustrates a simple graphical way of representing coded dynamics for the tent map by using the technique of finite directed graphs. Transition graph G is constructed using a covering rule: for example, an edge from vertex i to vertex j is drawn if and only if the image $f(I_i)$ of I_i contains the interval I_j. Coded dynamics is then described by all possible walks through the labeled directed graph G. And in fact, if we have been able to construct a covering partition P of the domain of the corresponding dynamical system, its associated coded dynamics can be effectively represented by a graph.

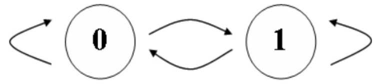

Figure 1: transition graph G representing the coded dynamics of the tent map with the partition $P = \{I_0, I_1\}$.

As a variation of the tent map, let f be the unimodal function defined by

$$f(x) = \begin{cases} 2x + \frac{1}{3} \text{ for } 0 \le x \le \frac{1}{3} \\ \frac{3}{2}(1 - x) \text{ for } \frac{1}{3} \le x \le 1 \end{cases}$$

where the partition $P = \{I_0, I_1\}$ is now I_0 = [0, 1/3] and I_1 = [1/3, 1], with the graph of f no longer symmetric with respect to the inversion point and thus the two branches no longer mirror-images of one another[7]. In this case it can be shown that the representation of the corresponding coded dynamics is given by the transition graph of figure 2.

[5] See, for example, Benhabib and Day (1980, 1981, 1982), Boldrin and Montrucchio (1986), Day (1982, 1983), Day and Shafer (1985) and Hahn (1992) for some economic models with a final dynamic unimodal equation. See also Sordi (1992) for a review on chaotic unidimensional models in macroeconomics.

[6] Moreover, the symmetry between expansion and contraction phases hid the fact that their basic difference was of a qualitative nature, and prevented their real understanding (see Goodwin, (1951)).

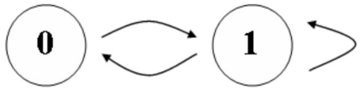

Figure 2: transition graph representing the coded dynamics of the modification of the tent map with the partition $P = \{I_0, I_1\}$.

It is not hard to see that the set of symbolic trajectories is exactly the set of binary sequences that cannot contain strings with two or more $0's$ in sequence, while the previous version admits any sort of infinite binary sequences, restricting the shift map to this new partition, therefore, fully describes the possibilities of the two-regime dynamics it can generate, by excluding one (sometimes important) dynamic possibility. One can see it as a way to fine-tune a posited canonical model, the tent map in this case, to fit certain data, and it is a good teaching example for a general approach to model recovery or identification (in an extension of the econometric language) to which we will come back later in the paper. This is the promising way ahead for coded analysis as applied to empirical socioeconomic data.

In the previous examples, we have used a *Markov partition* of the system domain, i.e., a partition yielding an almost one-to-one relationship between trajectories and symbolic sequences. If we use a partition that is not Markov, the description in the coded dynamics possibly it cannot be represented via transition graphs.[8] But as long as we can make use of one such partition, life is relatively easier, and we can still study the fundamental properties of the original model in the relatively simpler setting of coded dynamics. One of those key properties is *topological entropy*, one of the premier numerical invariants of a dynamical system serving many a purpose. From the dynamic viewpoint, entropy measures complexity and randomness; from the information-theory viewpoint, on the other hand, it is a measure of information capacity or ability to transmit messages. In fact, the number $\# B_n(X)$ of n blocks appearing on points of a space X of symbolic sequences gives us an idea of the complexity of X (the greater such a number, the more complicated the space). Instead of using the individual numbers $\# B_n(X)$ for $n = 1, 2, \ldots$, we can summarize their *collective* behavior by computing their growth rate as n varies. As can be readily seen, $\# B_n(X)$ grows approximately at the rate $2cn$, with the constant c as the growth rate. Then, for large n such value is approximated by $(1/n) \log_2 (\# B_n(X))$. This motivates the definition of entropy, to recall:

Definition *Let X be a space of symbolic sequences. The (topological) entropy of X is defined by*

$$h(X) = \lim_{n \to +\infty} \frac{1}{n} \log_2 (\# B_n(X))$$

With reference to the previous pedagogical examples, note that if X is a space of sequences defined in a set of R symbols, then $\# B_n(X) \leq R^n$ for all n and this implies that $h(X) \leq \log_2 R$. If X is the space of all symbolic sequences of R symbols, we have that $\# B_n(X) \leq \# R^n$, so it is $h(X) \leq \log_2 R$. Then, the two-regime dynamics of the tent map has entropy $\log_2 2 = 1$. On the other hand, for the modified tent map, we know that the set of feasible coded sequences is a set of binary sequences that does not contain the string 00. In this case it can be shown that the entropy of the regime dynamics is $\log_2\left(\frac{1+\sqrt{5}}{2}\right)$[9]. But life is hardly this easy, as Markov (endogenous) partitions, have no reason to coincide with exogenous partitions that are introduced on the basis of intuition or criteria for testing existing theories. The two obey two distinct necessities, and each of them places restrictions and demands upon the sought result the other does not care for (and, in general, covering partitions are seen as too stringent to applied sciences). Whenever, we have to make do without them, we can still either dive into data and resort to statistical methods (next section) and/or construct computational experiments (see the last section).

Data analysis of symbolic time series

Symbolic Time Series Analysis (STSA) is a statistical approach that arises in a natural way in the context of data analysis and processing of multiregime phenomena. We start with a given set of measurements which in some applications is a set of vectors of time series data representing the evolution of a dynamical process in its state space D. If the latter is endowed with a suitable (endogenous or exogenous) partition so that a menu of dynamical regimes can be identified,

[7] As was in business cycle theory, see Goodwin (1951) again.

[8] See Bolt et als. (2001) for a detailed study of non-generating partitions in one dimensional dynamics.

[9] See Brida and Punzo (2003) and references therein for a detailed exposition of representation of symbolic dynamics via directed graphs and for the techniques to compute the entropy.

coding them synthesizes the evolution of the *observed* system into a symbolic sequence, each symbol being a dated regime state. However, this coding approach and the related techniques can be easily extended to cope with other socioeconomic research settings where a finite number of states, at times identified by qualitative rather than quantitative characteristics, can be identified and therefore can be *tagged* accordingly. Thus, assume a sequence of data $\{x_1, x_2, \ldots, x_t, \ldots, x_T\}$ made up of vectors $x_t \in D \subset R^q$, for $t = 1, 2, \ldots, T$. Then R^q is the state space with the coordinates of the process under observations, subset D is the set of admissible values for such coordinates, a given dated vector x_t is the system state at time t in our choice of coordinate axes and, finally, q will be the number of state variables of the given dynamical system on D. Suppose that D is endowed with a partition of regimes. Then, we transform the sequence of data $\{x_1, x_2, \ldots, x_t, \ldots, x_T\}$ into the sequence of symbols $s_1s_2\ldots s_t\ldots s_T$, where $s_t = s$ if and only if x_t belongs to the regime labeled by s[10]. Once encoded, regime dynamics of the multiregime system can be described:

i) by the pattern of symbols within the associated string: i.e., which symbols appear and which do not (in other words, which regimes are visited and which not; and by their connectedness, i.e., how regions are visited in sequence, one after the other);

ii) by the normalized time spent in each region, or length of subsequences in the string exhibiting the same symbol;

iii) by the various cycles of different lengths (or periods) where a sub-set of symbols repeats itself in a structured sequence;

iv) finally, by the eventual emergence of particular patterns of symbols (structural cycles as asymptotic behaviors) and, perhaps, the disappearance of other cycles (transients).

Taking a population of strings, one can compute their statistics: e.g.

i) the frequency of visiting certain regimes;

ii) the frequency of paths connecting selected regimes, one-way paths and/or cycles (or closed loops, as against open loops);

iii) the average time spent in a given regime, path, cycle and the like.

STSA addresses the issue of how extract and describe time patterns of complex dynamical processes and recover information about the class of generating models. Compared with more standard analytical approaches, the novelty is in that this set of method does not rest upon any hypotheses as to the data generating (deterministic and/or stochastic) model. STSA accepts the *paradigm of irregularity*, where irregularity is the generic, and regularity the rare property of time series reflecting a dynamical process endowed with sufficient complexity (as social and economics processes redeemed to be). Under such paradigm, we have to look at most for near-regularities in the raw data sets, or at higher levels of dynamics of regime or coded dynamics.

The next step in identification of temporal patterns is the extraction of *short symbol sequences* of chosen length, from the overall sequence of symbols $s_1s_2\ldots s_t\ldots s_T$ coding the whole history of a system in terms of regimes, and we do this by grouping symbols together while preserving their temporal order. Such ordered subsequences are called *words* in the symbolic dynamics literature. They stand to us for *paths* (to adhere to dynamic terminology) or *patterns* to emphasize their qualitative significance looking at them from the regime point of view, while at times they will be called *episodes* to stress their occasional emergence from within an apparently unstructured symbolic string or erratic history. One may give a ready interpretation to some of such patterns, at least: e.g., any length-k sequence repeating the same symbol represents a sort of temporary equilibrium of the same duration but in a regime sense, thus allowing the system to wander around within a given regime domain (therefore, one such *temporary equilibrium* may be consistent with a broad variety of dynamics *within* the regime's domain). On the other hand, we may have fluctuations across regimes, some of them may show some periodicity, as they are in a conventional business or economic

[10] This process of transformation of data into a symbolic sequence is called symbolization in the STSA literature and can be done in several ways. See Piccardi (2004), Hirata et als. (2004), Daw et als. (2003) and Bolt et als. (2001) for a discussion of symbolization methods and the relevance of selecting a good partition. In Hirata et als. (2004) some methods to approximate generating partitions from data are presented.

cycles, but most often such periodicity does not show up, and is being dominated by the emergence of partially structured patterns, where, e.g., regimes follow for sometime a known sequence. It is clear that whenever we have point equilibria and in particular (local or global) attractors, closed orbits and limit cycles of the classical theory, they turn up to be special cases of our MRD, where a whole regime domain or a sequence of visited domains are collapsed into single member points. As we will see, the advantage of our coding technique however is in that, while comprising those as special cases, it also permits to handle a whole variety of irregular behaviour and often still to find some *order* (in addition or in substitution to invoking unexplained stochastic shocks) at the same time as not having to place restrictions on the size of the data sets (where the regular behaviours of mathematical dynamics are deemed to be quite rare). Both cases are typical in the applied social sciences, and otherwise addressed with a variety of *ad hoc* methods. To extract information encoded in the strings so constructed, we can do many things, provided the strings are long enough (in terms of time observations or *history length*) and/or large enough (as measured by the number of parallel observation of similar systems at the same time, or *histories* of equal length). Either way, we may aim at, e.g., constructing a *symbol tree*. This is a graphical representation of the symbol statistics in a given coded history as a function of the length of patterns in what has been called the *available dynamic menu*. We compute the relative frequency of occurrence of all symbol sequences of length k in a system's symbolic history and, varying the length $k \leq 1$, represent them as a tree, one branch for each value of length k. Hence, the first level shows the probabilities of occurrence of the individual symbols or regimes (as patterns of unitary length), the second the probabilities of occurrence of paths with two (different *or* equal) symbols, and so on so forth. The symbol tree is a compact information summary of the regime dynamics under observation. Looking back at the two graphs in the previous section, one sees that the tree is assigning probability weights to each directed arrow (level 1) and connected arcs (for $k \leq 1$). The interpretation is interesting. At each node or level of the tree, which represents a stage in its history, a given system (if probabilities are derived from its long history) or an average system (if they are derived across histories of various systems) has a probability of staying put where it is, or else to move, and if it moves, this probability refers also to where to move. Such probabilities depend upon past history, the portfolio or menu of available alternatives and finally (when many are simultaneously present) upon the other systems have and are being doing. The width of this probability window is important and reflects the degree of freedom left given those conditioning elements. It is as if at each stage a system were to make a constrained choice about its future, the next step, but as this will influence to an extent those further down the road, it is deciding upon its whole future, without knowing how much until it has traveled sufficiently further down. Each branch of the symbol tree can be represented by a histogram. This construction yields a direct visual representation of the temporal structure in the observed data. The histogram depends upon the chose symbol sequence length k and generally there are no clear-cut, theoretical rules to determine which value of k are sufficient for detecting significant patterns, if they exist at all. Just like the selection of the partition, the selection of k can be carried out in several ways and optimal choices partly depend on the way the partition is chosen. Later, we will have to come back to this issue. Symbol sequence histograms are compact representations of the general dynamics embedded in given time series and can be used to compare data sets, whether for different systems or for different phases in one's evolution. On the basis of computed frequencies, one can discriminate histories with distributions fairly uniform across the partitioned set of regimes from more eschewed ones. In particular, given two regime histories with the corresponding coded sequences, we can evaluate how different they are by means of some measure of their distance. The most commonly used of such measures is, of course, the Euclidean distance from the k-th levels of their trees or, in other words, for k-long episodes, here re-defined as

$$E_{AB}(k) = \sqrt{\sum_i (A_i - B_i)^2}$$

Where A_i and B_i are the probabilities for the possible sequence code or *episode i* in the distinct ***k***-histograms *A* and *B* of the two systems[11]. Descriptions of other measures of distances are given by Keller *et al*. (2004) and Daw *et al*. (2003) and in references included in these papers. The Euclidean distance works like a metric in the space of all possible sequences providing a measure of the distance between different ***k***-histograms in terms of the probability of exhibiting like episodes: a greater distance implies that the dynamics in the two data set is very different. In the next section, the Euclidean distance will be used as a target function for fitting the parameters of a multiregime model. Finally, E_{AB} can also be useful in testing for time reversibility, commonly used as an indirect test to decide which category of models is more appropriate to describe a given data set.[12]

A traditional index for characterizing the inner structure of a symbol sequence is the Shannon entropy. This information measure implicitly gives the average degree of *organization* of a symbol sequence histogram at any sequence length ***k***. The Shannon entropy H(***k***) at the ***k***-th level of a tree is based on the probability distribution of sequences with length ***k*** in the symbol sequence

$$H(k) = -\sum_i p_i \log_2 p_i$$

Here p_i is the probability of finding the *i*-th sequence of length ***k***, approximated by the number of times that the sequence indexed by *i* can be found in the coded history divided by the number of all such ***k***-long sequences. The base-2 logarithm expresses entropy in units of bits, so that *H*(***k***) is simple to interpret: it measures the average number of bits needed to specify an arbitrary sequence of length ***k*** in a symbolic sequence. It is a measure of the complexity of the data-generating process, or of observational variety: possible paths that still are not observed do not enter the measure as they get zero probability weight and the latter can be interpreted as an index of uniformity analogous to the standard deviation[13]. Because of the finite length of the data, many possible sequences may appear not to be realized which has been shown to significantly affect entropy estimates. *Normalized entropy* avoids this problem

$$H_S(k) = -\frac{1}{\log_2 N} \sum_i p_i \log_2 p_i$$

where *N* is the total number of *observed* episodes of length ***k*** (i.e. the number of sequences of length ***k*** with non-zero frequency). Clearly $H_s(k)$ is a modified form of *H* (***k***) and can be interpreted as the average uncertainty per symbol (or sequence). A simple calculation shows that, whatever ***k***, it is $0 \leq H_s \leq 1$ with $H_s = 1$ if and only if p_1, …, p_k, while $H_s = 0$ if and only if $p_i = 1$ for some *i*. In other words, modified entropy is zero when observational variety is absent and maximal when all the ***k***-histories have the same number of observations. *In this sense, entropy can measure the degree of randomness embedded in a given history*.

When working with equiprobable partitions, entropy may be useful to sort out, whether dynamics involved is deterministic or stochastic. In fact, each symbol sequence of length ***k*** must be equiprobable in truly random and sufficiently large set of data. In this case, $H_s = 1$ while $0 \leq H_s < 1$ for non-random data, lower values of H_s implying more deterministic structure. *A significant deviation from equiprobability is evidence of time dependence and deterministic structure in the data*[14]. With a preassigned partition, the value of the normalized entropy is a function of ***k*** and typically H_s decreases monotonically as ***k*** is increased from 1. Its global minimum H^* shows the value of ***k*** and thus the set of symbol-sequences which best distinguishing data from a random sequence. Symbol sequences that are *too short*, loose some of the important deterministic information while symbol sequences that are *too long*, basically reflect noise and data depletion. Then, *the value of the symbol sequence length **k** that minimizes H_s (**k**) can be seen as an optimal choice for the given data and selected partition.*

Model reconstruction for MRD

One powerful application of STSA is to contrast simulated data obtained from an assumed model, with real data and check whether differences are statistically significant. In this section, we consider the inverse problem of reconstructing a multiregime dynamical model from measured time series data. The problem is

[11] It means that we have to take the difference between corresponding probability entries at the given level of the two trees. Like the symbol sequence histograms, is a function of the symbol sequence length .

[12] For applications of time reversibility in symbolic time series see Daw et al. (2003).

[13] The average is computed only for categories of observations that effectively occur and this is reflected by the convention.

[14] Characterization of symbolic sequences is not restricted to the estimate of Shannon entropies. A broad range of so-called measures of complexity allows a more detailed characterization of the structure of symbolic sequences. See Kurths et al. (1996) for a good discusion of complexity measures.

framed as follows: a multiregime dynamical model is defined on a state space S that depends on a vector of parameters λ and there is an associated set of observed data to be emulated. That is, $S \subset R^m$ is the state space, $f_\lambda : S \rightarrow S$ is a parametric dynamical rule defining the model and $\{x_1, x_2, \ldots, x_t, \ldots, x_T\}$ is the observed data set. We assume that the model has n distinct dynamical regimes associated with domains $S_1, \ldots, S_n$ and we are interested in computing the parameter vector λ such that the model reproduces the observed regime dynamics. That is, we want to find the values of vector λ such that the model produce by iterations sequences that, once symbolized, are close *enough* to the coded string of observed dynamics. In other words, our goal is to find a model that generates a symbol tree similar according to some criterion to the observed one, all the way down to some predetermined level k. In general, recovering a model from a time series is done by varying the model parameters such that some error function is minimized. The usual methods to fit models using minimum squares methods adjust their parameters in such a way that the point sequences generated by iterating the posited model and observed data series appear similar. In general, the solution will not well reproduce the actual regime dynamics, and we will have to adjust the model in order to improve it. In order to do this, we will use statistics for the observed symbol sequence as a target for measuring the goodness of fit of the proposed model, because this statistics was shown to synthesize regime dynamics. This motivates the introduction of a new error function to be minimized, the function measuring the distance between observed and artificial regime dynamics. Our error function is the Euclidean distance between levels of the trees described earlier. This distance depending on the symbol sequence length k, we have to select a value of k and this can be done by minimizing the modified entropy of the observed symbolic sequence.

Let us describe the method step by step. The first step is to code observed data $\{x_1, x_2, \ldots, x_t, \ldots, x_T\}$ into the sequence of symbols $s_1, s_2, \ldots, s_t, \ldots, s_T$, and construct the observed symbol tree T_0. On the other hand, we iterate the initial condition x_1 using the function f_λ to produce an artificial time series $\{y_1, y_2, \ldots, y_t, \ldots, y_T\}$ where $x_1 = y_1$ and $y_{t+1} = y_t$ for $t = 1, \ldots, T-1$. At this point we can code $\{y_1, y_2, \ldots, y_t, \ldots, y_T\}$ into $\{t_1, t_2, \ldots, t_t, \ldots, t_T\}$, where $t_t = i$ if and only if y_t belongs to the regime labeled by S_i. From this sequence, the artificial symbol tree T_A is constructed.

Next step is to compute the value of $\hat{k}$ that minimizes the modified entropy of $s_1, s_2, \ldots, s_t, \ldots, s_T$. Then, the error function $\mathcal{F}(\lambda)$ is the Euclidean distance between the branch vectors of T_0 and T_A at level $\hat{k}$

$$\mathcal{F}(\lambda) = E_{T_0 T_A}(\hat{k}) = \sqrt{\sum_i (T_{0i} - T_{Ai})^2}$$

The error function $\mathcal{F}(\lambda)$ is defined over the parameter space and its plot represents the error landscape. Reconstructing the generating dynamic model requires to find the global minimum in this landscape. The complication is that this search is no elementary task, as we have no general formula for the error function $\mathcal{F}(\lambda)$ and we are only able to compute the value of $\mathcal{F}(\lambda)$ for given vector λ. Therefore, as one do not know where to look for a solution or where to start from, most conventional search routines will have difficulty too in locating the sought solution. In Brida and Garrido (2006), it is shown that a genetic algorithm as the routine to find the optimal values of the parameters appears to perform relatively well (even for a chaotic map) and is capable of finding a good approximation for the global minimum. Tang *et als*. (1994) use the simulated annealing algorithm, applying symbolic methods to reconstruct the Hénon and Ikeda maps and showing that the method is highly robust even in presence of observational and dynamical noise.

Summary and conclusions

This paper makes a move towards unifying applied and theoretical analyses of complex economic dynamics by introducing multiregime models and some new tools of qualitative analysis of data in the presence of multiple regimes. We recall that qualitative analysis of data produced by quantitatively capable models has an established tradition in economic dynamics and partly in the theory of dynamical systems, from which economists borrowed it. In the latter context one resorts to qualitative analysis whenever closed forms ensuring explicit solutions cannot be obtained, deploying a variety of methods, basically of topological nature, which permit a limited understanding of some general features.

The dynamical construction based upon the notion of regime: i) reproduces state indeterminacy typical of complex behaviors; ii) it does away with the juxtaposition between deterministic and stochastic interpretations of such indeterminacy; finally, iii) its value added is in handling large coordinate systems. Effectively reducing the space dimension or the number of variables to be handled, it is a means to summarize information for large sets. One recognizes that the twofold notion of a regime and of a whole set of them as a portfolio of qualitatively different dynamical behaviors, does play a central role in many theoretical settings in economics and the social sciences. It often seems more natural than the notion of a dynamical state, as is defined in classical physics and currently used both in the theoretical dynamics and in dynamic econometrics. Economic modeling is concerned, almost always, with whole vectors of values of economically relevant coordinate variables that are required in a certain sense to conform to some implicit consistency requirement. This is the basic requirement for the definition of a regime over the set of assumed relevant variables. We do not see why this common sense cannot make room into the formal approach. Is it a matter of logics in the high fundaments or lack of mathematical tools? We surmise it is the fear of venturing deep into a land where known and easy methods do not help, and fresh ones have to be invented, not discovered in a mathematical book.

The theoretical exercise has therefore a natural empirical counterpart, often an empirical foundation. The connection is established via computational experiments. The standard computational experiment may go as follows. At step one, we introduces taxonomy of regimes for the problem at hand, this being determined by the theoretical beliefs or by the adoption of one or more theoretical frameworks. Then, actual data is processed in the statistical descriptive structure that is coherent with the theoretical choice. By analyzing the descriptive dynamics, the minimal number of partitions in the regimes space that is relevant is determined endogenously. Finally, one can estimate a model, which is subject to the constraint of reproducing the coded dynamics as closely as possible and use it for computational experiments. Its forward output is a qualitative prediction of a sequence of regimes, rather than the quantitative prediction of asymptotic states of standard econometric techniques manipulating cross section and/or time series data. Qualitative predictions are compared to actual histories, to check for the goodness of the pre-determined model. Such computational heuristics can be evaluated against other strategies available in the literature, on one side those in the Artificial Intelligence tradition that dominate the computational approach to economics; on the other, with those generated by the Real Business Cycles approach. Qualitative analysis is re-defined as a set of procedures (some of them formal, some merely heuristic) for pattern recognition and related concepts, rather than being a mere alternative to analytical dynamic analysis. Our approach focuses, therefore, upon identifying recognizable patterns embedded into economic time series whose variety and irregularity could otherwise be entirely attributed to stochastic components. Thus, it can be seen as an experiment with a deterministic viewpoint where adjustment and structural dynamics are not (and intentionally cannot be) distinguished from one another and therefore are simultaneously dealt with. The extension of our approach to qualitative analysis to fields in the social sciences where this has always played a dominant role seems natural.

Acknowledgements: This research was carried on with the financial support of the Free University of Bolzano, project "Dynamical Regimens in Economics modeling and statistical tools".

References

ALLIGOOD, K. T.; SAUER, T. D.; YORKE, J. A. **Chaos:** An Introduction to Dynamical Systems. Springer-Verlag, New York, 1997.

BENHABIB, J.; DAY, R. H. **Erratic accumulation.** Economic Letters, 6, 1980, p. 113-117.

BENHABIB, J.; DAY, R. H. **Rational choice and erratic behaviour.** Review of Economic Studies 48, 1981, p. 459-471.

BENHABIB, J.; DAY, R. H. A characterization of erratic dynamics in the overlapping generations models. **Journal of Economic Dynamics and Control.** 4, 1982, p.37-55.

BOLDRIN, M.; MONTRUCCHIO, L. On the indeterminacy of capital accumulations path. **Journal of Economic Theory.** 40, 1986, p. 26-39.

BÖHM, B.; PUNZO, L. F. Detecting Structural Change: A Scheme for the Comparison of Austria and Italy in the Seventies and Eigthies. In: CLAUSER, O.; KALMBACH, P.; PEGORETTI, G.; SEGNANA, L. (Eds.). **Technological Innovation, Competitiveness, and Economic Growth.** Duncker & Humbolt, Berlin, 1992.

BÖHM, B.; GAFFARD, J. L.; PUNZO, L. F. Industrial dynamics and employment in Europe. **Final Research Report to the European Commission**. TSER Programme, 2001.

BOLLT, E. M.; STANFORD, T.; LAI, Y.-C.; ZYCZKOWSKI, Y.-C. What symbolic dynamics do we get with a misplaced partition? On the validity of threshold crossings analysis of chaotic time-series. **Physica.** D 154, 2001, p. 259-286.

BIMONTE, S.; BRIDA, J. G.; PUNZO, L. F. **Notions of Regime**: a review. I.D.E.E. Working Papers Series 2001-02, 2001.

BRIDA, J. G.; PUCHET ANYUL, M.; PUNZO, L. F. Coding economic dynamics to represent regime dynamics: a teach-yourself exercise. **Structural Change and Economic Dynamics.** 14, 2003, p. 133-157.

BRIDA, J. G.; PUNZO, L. F. Symbolic Time Series Analysis and Dynamic Regimes. **Structural Change and Economic Dynamics.** 14, 2003, p. 159-183.

BRIDA, J. G.; GARRIDO, N. Exploring two inflationary regimes in Latin-American economies: a Binary Time Series Analysis. **International Journal of Modern Physics C**. 17, 2006, issue 1.

DAW, C. S.; FINNEY, C. E. A.; TRACY, E. R. A review of symbolic analysis of experimental data. **Review of Scientific Instruments.** 74, 2003, p. 916-930.

DAY, R. H. **Complex Economic Dynamics**. Vol. 1 and 2. The MIT Press, Cambridge, Mass, 1994, 2000.

DAY, R. H. Irregular growth cycles. **American Economic Review.** 72, 1982, p. 406-414.

DAY, R. H. The emergence of chaos from classical economic growth. **Quarterly Journal of Economics.** 98, 1983, p. 201-213.

DAY, R. H.; SHAFER, W. Keynesian chaos. **Journal of Macroeconomics.** 7, 1985, p. 277-295.

DURLAUF, S. N.; QUAH, D. The new empirics of economic growth. In: **Handbook of macroeconomics**. Volume 1A. Amsterdam, New York and Oxford: Elsevier Science, North-Holland, 1999, p. 235-308.

GOODWIN, R. M. The Non-Linear Accelerator and the Persistence of Business Cycles. **Econometrica.** 19, 1951, p. 1-17.

HAHN, F. R. **Equilibrium dynamics and chaos**: A textbook approach. Jahrbücher für Nationalökonomie und Statistik, 210, 1992, p. 279-285.

HIRATA, Y.; JUDD, K.; KILAMINSTER, D. Estimating a generating partition from observed time series: symbolic shadowing. **Physical Review E.** 70, 2004, 016215.

KELLER, K.; WITTFELD, K. Distances of time series components by means of symbolic dynamics. **International Journal of Bifurcation and Chaos.** 14, N. 2, 2004, p. 693-703.

KURTHS, J.; SCHWARZ, U.; WITT, A.; KRAMPE, R. Th.; ABEL, M. Measures of complexity in signal analysis. In: KRATZ, R. A. (Ed.). **Chaotic, Fractal, and Nonlinear Signal Processing**. AIP Press Woodbury, New York, 1996, p. 33-54.

PICCARDI, C. **On the control of chaotic systems via symbolic time series analysis**. Chaos 14(4), 2004, p. 1026-1034.

TANG, X. Z.; TRACY, E. R.; BOOZER, A. D.; DEBROWN, A.; BROWN, R. **Reconstruction of Chaotic Signal Using Symbolic Data**. Physical Letters A 190, 1994, p. 393-398.

SORDI, S. Chaos in macrodynamics: an excursion through the literature. **Quaderni del Dipartimento di Economia Politica**. N. 195, Siena, 1996.