Time Abstraction and Quantitative/Qualitative Interpretation of Multiple Dynamics Processes

Laurent Ayrolles, Robert Faivre and François Guerrin

Unité de Biométrie et d'Intelligence Artificielle, INRA
BP 27, 31326 Castanet-Tolosan Cedex France
Phone: +33 61 28 53 34; Fax: +33 61 28 53 35
E-mail: ayrolles@toulouse.inra.fr

Abstract: This paper describes a qualitative reasoning approach aimed at representing and interpreting a dynamic process evolution. The problem specifically addressed is the presence of multiple time-scales in complex systems. Definitions of a temporal granularity as well as related concepts are provided. For the representation of a single process, a segmentation and abstraction method is described. The identification of dynamic features at any level of abstraction then supplies an help to better choose relevant sampling frequencies of the simulated process. Examples are given for a crop growth model and a prey-predator model. Interpretations are confronted to expert knowledge.

1. Introduction

The coexistence of processes operating at different time-scales is frequent in complex systems. We can informally define for a dynamic process the time-scale as 'the minimal duration in which any significant change may be detected'. Analysing the behaviour of a system containing such processes arises several issues. 'Significant change' is somewhat ambiguous. It supposes several time-scales might fit and so different representations of a same process might be relevant for some purposes. One task is therefore to help the user in choosing relevant time-scales to sample simulation outputs or to set a measure apparatus. This relevancy not only depends on the user's needs but also on the process intrinsic fluctuations. A second task is to help the user to confirm or invalidate his choices. This requires to be able to describe and interpret each process at several time-scales.

In some fields of applications, e.g. ecological systems, precise results are often of poor meaning whereas the shape of an evolution is far more instructive. However explicit time is needed. For these reasons we chose a quantitative/qualitative approach to interpret time evolution curves (principally simulation outputs).

Many papers deal with time abstraction and granularity notions, defining and using them in various ways [1, 2, 3, 4, 5]. We emphasize here three Qualitative Reasoning approaches [6, 7, 8] representative of the time-scale notion although addressing the problem of prediction which is not ours; it is worth noticing that two of them refer to biological systems [6, 8].

In QSIM, Kuipers [6] forms a hierarchy of system variables ('faster' and 'slower') according to the time they need to get back to the equilibrium state after a perturbation. Iwasaki [7] bases the representation on a hierarchy of distinct models. Each of them is characterized by its 'temporal granularity' T_M, defined as follows: if $T_M = n$, then any duration shorter than 10^n seconds is taken as instantaneous. From user-defined simulation horizon and temporal granularities of the variables of interest, the corresponding model is generated. Rickel and Porter [8] first specify temporal scales for a reduced set of 'reference' variables. Their system then determines for the other variables consistent time-scales to answer user's queries about the behaviour of a variable of interest. The processes which can be considered as instantaneous are isolated from those to be simulated.

In these three approaches, time-scales (or granularities) are arbitrarily defined orders of magnitude. Apart from their lack of objectivity and flexibility, these definitions do not fit processes that are not known to be 'achieved' after a certain amount of time (i.e. reaching an equilibrium, completing a cycle, etc.). Moreover, no help is supplied to set or adapt scales. This may be more embarrassing, whether the aim is prediction or interpretation.

We propose in section 2 a definition of temporal granularity to overcome the flexibility drawback. In section 3 we describe a procedure of process segmentation based on Cheung and Stephanopoulos' works [9, 10]. An evolution is decomposed into geometric episodes independent of any user-defined 'scale of interest'.

Each episode is related to a particular qualitative shape of behaviour, and carries both symbolic and numerical information. An algorithm aiming at successively abstract such a segmented representation, from the finest to the coarsest level, is also described.
Numeric, symbolic and statistic tools of interpretation are proposed in section 4. They allow one to point out some standard dynamic features at a given level of abstraction, as well as to choose relevant time-scales. Finally (section 5), examples of interpretation of simulation outputs in agronomy and ecology are provided, checked against available expertise.

2. Basic concepts and definitions

Setting *a priori* a time-scales hierarchy can hide intermediate interesting phenomena. In order to avoid such arbitrary choices, context-free definitions are needed. We start from actual data, characterized by their temporal range and the time-interval between them. Let us define:
— a *system* as a set of variables X of which evolutions are observed or simulated.
— a *qualitative state* of a variable X at the time-point t_i $(i = 0, \ldots, n)$ as the triple:
$S(X, t_i) = < x_i, dir(x_i), ddir(x_i) >$, where x_i is the numerical value, $dir(x_i)$ the qualitative tendency and $ddir(x_i)$ the qualitative curvature of X at $t = t_i$. Tendency and curvature take their values on the four symbolic elements set $\{inc, std, dec, ?\}$, corresponding to increasing, steady, decreasing, and unknown evolution, respectively. The system state is the set of every variable states.
— a *process* as the sequence of states of X over $t_0, \ldots, t_n$: $X_{t_0, \ldots, t_n} = \{S(X,t_0), \ldots, S(X,t_n)\}$.
— a *temporal domain* $\Delta = [t_0, t_n]$ over which the process is observed or simulated.
— a *time-interval* as any subdivision of the domain: $\delta_{ij} = [t_i, t_j[$ such that $t_i < t_j$.
— the *temporal granularity* of a time-interval with respect to the temporal domain as:
$G(\delta_{ij}) = |\delta_{ij}| / |\Delta|$, where $|\cdot|$ stands for durations.
We have $0 < G(\delta_{ij}) \leq 1$, where 1 corresponds to the maximal abstraction ($\delta_{ij} = \Delta$). A process is represented over a sequence of time-intervals; their granularities are a good way to characterize the level of detail of the process representation. Abstraction corresponds to a granularity increase (removal of points) and refinement to a granularity decrease (adding new points). Granularity is a relative and dimensionless notion. Thus we think possible to deal with any process, whatever its particular domain and time-intervals may be. This notion has been discussed in [11] with reference to the representation of simulation outputs at distinct abstraction levels.

3. Triangular and trapezoidal representations

Processes do not necessarily fluctuate regularly over time. A sequence of variable and non-arbitrary granularities must thus describe them. Starting from a chronological series, i.e. a pointwise evolution, our aim, as in [9, 10], is to derive different representations as sequences of geometric episodes (triangles, trapeziums) able to support a behavioural interpretation.

3.1. Hypothesis and definitions

Let us assume that the process does not contain any slope discontinuity (class C1). We call *distinguished time-point* (*dtp*) a time-point t_i where the process exhibits either a local extremum or an inflexion-point; *dtp* t_i is such that:
(1) $dir(x_{i-1}) \neq dir(x_i)$ and $dir(x_{i-1}) \neq std$ and $dir(x_i) \neq std$ (extremum), or
(2) $ddir(x_{i-1}) \neq ddir(x_{i+1})$ and $ddir(x_{i-1}) \neq std$ and $ddir(x_{i+1}) \neq std$ (inflexion point).
The set of *dtp* of $X_{t_0, \ldots, t_n}$ is noted $DTP(X)$.

A *triangular episode* delimits the open time-interval existing between two consecutive *dtp*'s, i.e. $]t_i, t_k[$ is a triangular episode if:
$t_i, t_k \in DTP(X)$, $t_i < t_k$, and $\neg\exists\, t_j \in DTP(X)$ such that: $t_i < t_j < t_k$.
It therefore corresponds to a time interval on which the process keeps qualitatively the same behaviour, i.e. on which all the points have same tendency and curvature. We denote it $T(X, t_i)$ when it starts at t_i. Tendency at t_i, and tendency and curvature at t_{i+1} are sufficient to determine the eight possible shapes of triangles (see Table 1); a line (flat episode) is merged with adjacent portions of behaviour to form a single non-flat episode.
The four main types of triangles are named A, B, C and D: A is an accelerating increase, B a slowing down increase, C a slowing down decrease and D an accelerating decrease. The particular types A^*, B^*, C^* and D^* are distinguished as denoting an equilibrium point.

dir (x_{i+1}) → / *ddir* (x_{i+1}) ↓	*dec*	*std*	*inc*
dec	D D^*	B^*	B
std			
inc	C	C^*	A A^*

Table 1: The eleven shapes of triangular episodes according to tendency at t_i and tendency and curvature at t_{i+1}. The distinction between D and D^*, and A and A^* is made according to the value of $dir(x_i)$.

Because of the continuity assumption, the following sequences may only be found: AB, BA, CD, DC, BD and CA, as well as any combination with particular types (A^*B, B^*D^*, etc.). The four first pairs are made of two episodes with same tendency, and embody one inflexion point. The last two ones are composed of two episodes of same curvature and contain an extremum.

3.2. Triangular representation

We start from a series of points (x_i,t_i) corresponding to a process $X_{t_0,\ldots,t_n}$ (see Fig. 3 b). We assume its sampling frequency is regular, thus, for $\delta_i = [t_i,t_{i+1}[$, $G(\delta_i) = 1/n$.

Tendency and curvature are assigned to each time-point, using simple increments rules: sign of the difference between two values for tendency, sign of the relative difference between three values for curvature.

To obtain a representation fitting the process fluctuations, we split $X_{t_0,\ldots,t_n}$ into a series of triangular episodes, called *triangular representation* (see Fig. 3 b) and denoted:

$T(X_{t_0,\ldots,t_n}) = \{T(X,t_0), \ldots, T(X,t_m)\}$, where $t_0, \ldots, t_m$ are the *dtp*'s of the initial evolution.

The initial process is contained in the envelope of triangles. So each episode preserves a part of the original numerical information. An iterative algorithm, considering three successive points, performs abstractions and refinements according to cases:

— abstraction: (1) if several points are aligned ($ddir(x) = std$), then only the boundary ones are kept; (2) if several consecutive points have same tendency and same curvature, then they are abstracted as a triangular episode.

— refinement is performed: (1) when an inflexion point is detected between two time-points; (2) when a steady behaviour lasts at least one time-step between a change of tendency from *inc* to *dec* or *dec* to *inc*. An interpolation is then performed between two known points, and the granularity becomes locally finer.

3.3. Trapezoidal representations

To abstract X to any level, we first transform triangles into trapeziums. Then local abstraction of trapeziums is performed recursively.

A trapezoidal episode is the geometric combination of two consecutive triangular episodes. It thus represents an interval of same tendency *or* of same curvature, whereas a triangular episode is an interval of same tendency *and* of same curvature. There is no loss of qualitative behavioural information (i.e. tendency or curvature). According to the continuity assumption, the only possible associations of triangles are the following (see Fig. 1):

— CA and BD are episodes with constant curvature and a change of tendency;

— BA and CD are episodes with constant tendency and an inflexion point.

Allowed transitions between trapeziums are presented in figure 2. Several constraints are specified to settle the way trapeziums can be constructed: boundary points, tendencies and curvatures are kept, and all the points of initial triangles belong to the resulting trapezium.

Transforming pairs of successive triangles into trapeziums results in a series of trapezoidal episodes, i.e. the *trapezoidal representation* (see Fig. 3 b):

$Z(X_{t_0,\ldots,t_n}) = \{Z(X,t_0), \ldots, Z(X,t_l)\}$, where $Z(X,t_i)$ is a trapezoidal episode beginning at t_i. It must be noticed that t_0 is not necessarily the same as in the triangular representation, as neither A or D can begin a trapezium.

The successive abstractions of the trapezoidal representation are performed step by step. Each step consists in grouping together the trapezoidal episode(s) of minimal duration with its (their) two neighbours, to form new trapezium(s). This yields a new trapezoidal representation denoted $Z^s(X_{t_0,\ldots,t_n})$ (see Fig. 3 b).

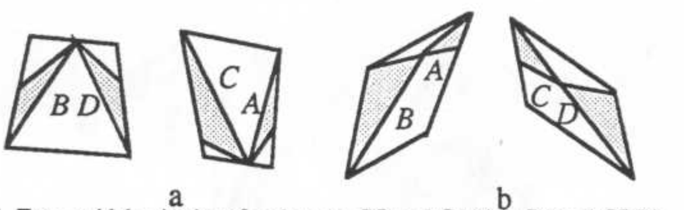

Fig. 1. Trapezoidal episodes of main types *BD* and *CA*(a) or *BA* and *CD*(b).

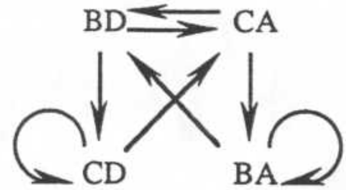

Fig. 2. The graph of possible transitions between trapezoidal episodes of main types.

The evolution at the step $s=0$ corresponds to the first trapezoidal representation. Each time three episodes are grouped into one, four tendencies and curvatures out of six disappear. Here occurs the first loss of behavioural information, but in an objective manner, i.e. without presuming the user's interest. The abstraction of three adjacent trapeziums is submitted to constraints similar to those needed to build initial trapeziums from triangles.

A new trapezoidal representation is not an envelope of the process anymore; it is impossible to keep the points of the three trapeziums while satisfying behavioural constraints. In fact, we chose to favour qualitative information rather than actual numerical values. A resulting trapezium must be understood as a time-interval for which the process characteristics (i.e. values, tendencies and curvatures) are only known in the neighbourhood of its boundary points.

4. Tools of interpretation

Segmentation and successive abstractions of evolution curves make easier a visual analysis. The granularity values allow a first characterization. Some other tools have been developed: structure of granularities, symbolic strings and spectra. The structure and some spectra were defined in [9, 10], but unfortunately without showing what automated reasoning could be based upon them.

4.1. Time-scale and fluctuation index

Granularities precisely measure *local* levels of detail: for the initial series of points, they take into account the duration between two consecutive time-points. In the triangular and trapezoidal representations, granularities are ratios between episode durations and the temporal domain. At any step of abstraction, the *minimal granularity* is used to set the *global* level of detail of the representation; we denote it G^s (where s is the step of abstraction). Granularity can be easily transformed into time duration. Particularly, G^s corresponds to the time-scale, as the *minimal duration* of a local variation.

Speed ('fast', 'slow') is a vague notion often employed when the context is precise enough to avoid ambiguities, namely when processes are known as periodic, reaching an equilibrium, etc., or are expressed using the same unit. Otherwise it is necessary to create a common unit: a *fluctuation index*, calculated over any interval $[t_i,t_j]$ as the ratio between the number of episodes in $[t_i,t_j]$ and its duration t_j-t_i; it is expressed in number of episodes per time unit.

4.2. Structure of granularities

The structure of granularities (see Fig. 3 a) is a diagram summarizing all the trapezoidal representations of a process. Boundary time-points of trapeziums of the first trapezoidal representation $Z^0(X_{t_0,\ \dots,\ t_n})$ are set on the x-axis. The y-axis bears the minimal granularity at each step of abstraction. For each step s, $Z^s(X_{t_0,\ \dots,\ t_n})$ represents the process at granularities belonging to $]G^{s-1},G^s]$. Moving upward in the diagram corresponds to abstraction, whereas moving downward is a refinement.

According to a minimal granularity specified by the user (denoted G), the structure of granularities provides indications about the relevancy of the sampling frequency of the process:

— if $G<G^0$, even the finest representation is too coarse; the sampling frequency must be increased;

— if $G>G^{last\ step}$, the last representation is valid; the one or two remaining episodes are enough to represent the process. However, by construction, our method is dependent on the boundary points of the process. We will discuss this point in section 6.

— in general, G is between two minimal granularities G^{s-1} and G^s; in this case, the system advises to sample the process with a frequency of G^{s-1}, and assesses the influence of this choice on the structure. The procedure is thus re-executed and a new structure plotted. The user performs a visual comparison between the two structures and the trapezoidal representations at G^{s-1}. If the similarity is sufficient for his purpose, the sampling frequency G^{s-1} reveals satisfactory, otherwise it is divided by two and the procedure re-executed. An example of this reasoning process is in section 5.

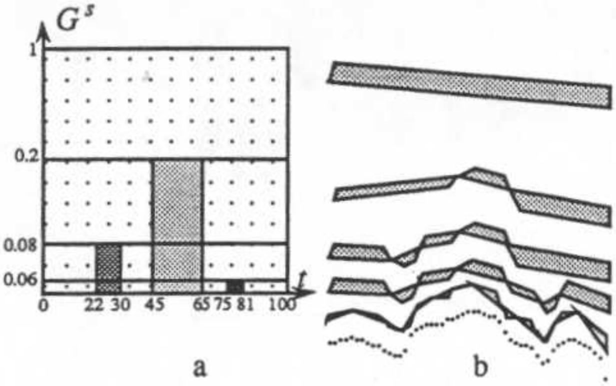

Fig. 3. (a) An example of structure of granularities. Over $\Delta = 100$, the process can be represented at 4 minimal granularities. (b) From bottom to top: the corresponding initial series of points, triangular and trapezoidal representations.

High-frequency fluctuations can be viewed as embedded in some low-frequency variations. In terms of fluctuation indices, it means that abstracted episodes belong to a highly fluctuating evolution, embedded in a low fluctuating one. The first is called *embedded evolution*, and the second *skeleton evolution*. Both are the components of the initial one. Getting them by decomposition supports the search for influences between processes, in terms of fluctuations.

This decomposition is worthwhile when a large gap between two consecutive minimal granularities appears in the structure: it highlights a feature of regularity, as it means many episodes with close durations are abstracted at the same step or after a few steps, on the same part of the temporal domain (see an example Fig. 10 c).

It is performed as follows when a block of three trapeziums is abstracted:

— the skeleton evolution (Fig. 4 b) is a trace of the global magnitude variation between the boundary points of this block. It is a new series of ordered points, that can be afterwards segmented and successively abstracted. It could be compared to a local coarse smoothing;

— the embedded evolution (Fig. 4 c) is the straightened initial evolution; we keep the abstracted episodes, but assign a zero value to the boundary points of each trapezium.

4.3. Symbolic strings

For any representation, the sequence of episodes types forms a symbolic string. Any kind of local response of a first or second order system is easily translated into triangular types, and thus easily detected in the triangular string (see Fig. 5 a). Analogously, the symbolic string of trapezoidal types allows one to identify the response of an oscillatory system (e.g. a resonant second order system; see Fig. 5 b). More behavioural features about oscillations are inferred (sinusoidal, damped, sustained, divergent, etc.) if we take into account additional information relative to magnitude variations and durations. Local stability is established when magnitude variations decrease in an oscillation.

To automate this, pattern matching is performed within symbolic strings, in order to detect particular behaviours or recurrent non-overlapping sequences. A periodic behaviour is identified when recurrent sequences are adjacent. Adding magnitude variations and durations may help discriminate (or conversely, unify) among recurrent sequences.

The initial set of possible sequences is formed by elementary patterns corresponding to two geometric episodes (i.e. 2 letters for the triangular string and 4 for trapezoidal ones): $\{AB, BA, CD, DC, BD, CA\}$ for the triangular string, $\{CABA,\ CABD,\ BDCA,\ BDCD,\ CDCA,\ CDCD,\ BABA,\ BABD\}$ for trapezoidal ones. An algorithm yields the set of recurrent sequences together with their location and number of occurrences: while the set of possible sequences is not empty, the places and number of occurrences of each pattern are searched; if a pattern occurs more than once, then two trapezoidal types are added at the end of it, to form two new patterns; otherwise it is withdrawn, etc.

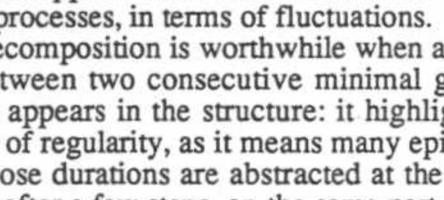

Fig. 4. Separation of the trapezoidal evolution (a) in the skeleton (b) and the embedded (c) evolutions.

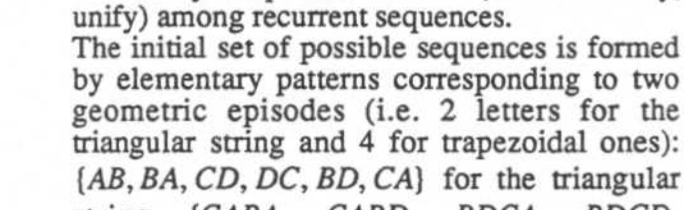

Fig. 5. Example of an oscillating behaviour. Its triangular (a) and trapezoidal (b) symbolic representation.

This simple algorithm is sufficient to assess recurrence in strings of hundreds of letters, since overlapping sequences are not searched and there are few possible transitions between episode types (see Fig. 2). Any episode addition and sequence withdrawal is kept in trees with initial patterns as roots. This execution trace allows the user to find out rapidly any specified sequence, corresponding to a unique pathway in the trees.

4.4. Spectra

The frequency spectrum of a temporal signal gives the distribution of frequencies over time. Analogously, specific tools are built to analyse the distribution of behaviours, at any single level of abstraction. Episodes have three meaningful features: type, duration and magnitude variation, set for the x-axis in three kinds of spectra (see Fig. 6). We will thus refer to: *types spectrum*, *granularities spectrum*, *magnitude variations spectrum*. The y-axis may stand for durations, magnitude variations or episodes count. Observed frequencies can be compared to a theoretical distribution using the χ^2 test.

Using granularities or magnitude variations spectra, one can exhibit classes (see Fig. 6 a and 6 c) and interpret such diagrams as usually performed (e.g. mode detection, relative percentage, etc.). Classes are interpreted as the various ways the process fluctuates (e.g. for two classes in a granularities spectrum, high versus low rate of change). Spectra may also be partial, namely addressing specific types of episodes. Using spectra thus leads to interpretation such as:

— *95% of BD are in the first class*, which indicates that maxima are (for the most) reached at the finest granularities;

— *BD and CA are the most frequent, CD and BA the least frequent*, which means that the behaviour is mainly oscillatory;

— *80% of the total magnitude variation is due to episodes BA*, which indicates that the process increases;

— *67% of the total number of episodes are in the last class*, which indicates that the main part of episodes have a large magnitude variation.

5. Examples of interpretation

5.1. A crop growth model: *EPIC (Erosion Productivity Impact Calculator)*

The aim of the *EPIC* simulator [12] is specifically to study the relationships between land farming and soil erosion. It describes all the processes relative to crop growth: biomass, evapotranspiration, leaf area index, soil capacity, photosynthesis, etc. Simulations are the only means for agronomists to forecast those dynamics. We analyse here the evolution of the actual evapotranspiration of wheat and maize in annual rotation, simulated during 4 years using a time-step of 1 day (see Fig. 7 a). This process (denoted *etr1*) is difficult to analyse visually, because of its erratic fluctuations. We applied first our methodology to simulation outputs, then the interpretations derived from it were checked in light of the knowledge of an agronomist.

Some features that cannot be distinguished within the initial evolution become clearer in our abstracted representations:

— At the granularity corresponding to the scale of 7 days, *plateaus* appear in the maize phases of growth (see Fig. 7 b). Their detection and temporal location were not so clear in the initial evolution.

— In the fourth cycle, the plateau begins and stops earlier than in the second cycle. This behaviour was related by the agronomist to a *major climatic event*, that we can date with a precision of one day.

— At the scale of 56 days, the only remaining cycles are those corresponding to wheat (see Fig 7 c). This confirms that *the cycle of growth is longer* for wheat than for maize.

The structure of granularities reveals that 29 steps are necessary to abstract entirely *etr1* (see Fig. 8 a). One gap is only observed at the end of

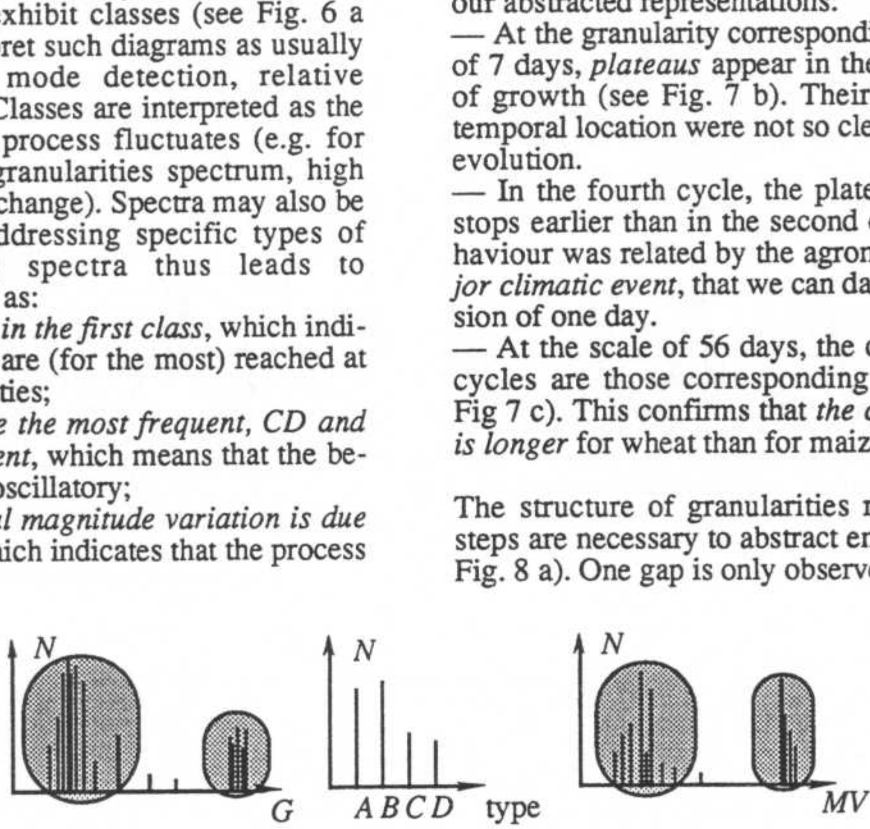

Fig. 6. Spectra expressed in number of episodes N: (a) granularities (G) spectrum. The tinted zones are classes of granularities; (b) types spectrum; (c) magnitude variations (MV) spectrum. The tinted zones are classes of magnitude variations

the abstraction, so the decomposition into the skeleton and embedded evolutions is not performed. To evaluate the influence of the initial sampling frequency upon the previous detection of plateaus (performed at the scale of 7 days), other initial series of points are tested: *etr7* (frequency: one point per week), and *etr30* (frequency: one point per month). On the corresponding structures of granularities (see Fig. 8 b and c), the representations at any minimal granularity appear very dissimilar. This is not surprising, as the process behaviour is erratic. A greater sampling frequency is thus suggested to minimize the influence of individual points: we take one point out of 3 days, i.e. approximately half a week (*etr3*). First, the new representation at the scale of 7 days exhibits the same plateaus as previously. Second, when we compare the new structure and the first one (see Fig. 8 d and a), similarities are found. Storing only one point out of 3 seems thus sufficient to retrieve the same plateau phenomena.

The pattern-matching procedure is applied to the symbolic strings obtained for processes *etr1*, *etr7* and *etr30*.
— In the representation of *etr1* at the scale of 7 days (see Fig. 7 b), a recurrence of {*BABABD CABDCABDCABDCABD*} appears in the first and fourth annual cycles. It corresponds to the presence of two peaks in each of these cycles. The second and the third cycles only have similar shapes for their increasing part (recurrence of {*CABDCABDCABDCA*}).
— If we take the representation of *etr7* at the same scale, that is 7 days, recurrent sequences are found in the first and the third cycles, and in the second and the fourth. This suggests a way to discriminate between the two species. However, each sequence represents approximately half a cycle only, which is not enough to confirm this assumption.
— Let us consider now the representation of *etr30* at its minimal scale, that is one month (see Fig. 7 d). The procedure detects two recurrent sequences {*ABDCABDCABDC*} corresponding to the entire cycles of wheat, and two sequences {*BDCABABDC*} corresponding to the cycles of maize. Hence, it is the right scale to observe inter-species differences, i.e. between maize and wheat behaviours, ignoring intra-species differences between cycles. The major difference between these behaviours is that a straight increase begins the growth of wheat (episode *CA*) whereas an inflexion-point slows down that of maize, causing a plateau (episode *BA*). This confirms the visual interpretation.

The types spectrum, expressed in number of episodes, derived from the first trapezoidal representation of *etr1* shows a large majority of *BD* (maxima) and *CA* (minima) episodes (see Fig. 9 a).

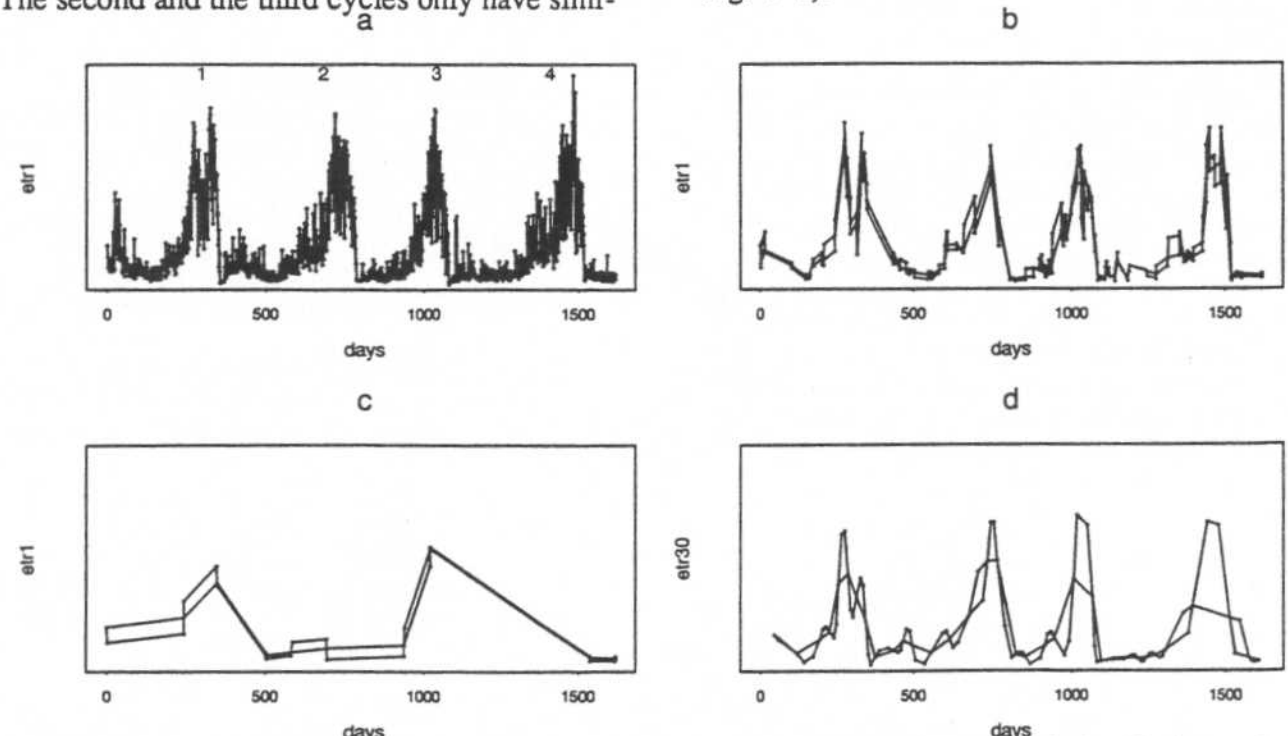

Fig. 7. Actual evapotranspiration of wheat and maize in annual rotation; (a) simulated evolution *etr1* (peaks 1 and 3: wheat; peaks 2 and 4: maize); (b) representation of *etr1* at the scale of 7 days; (c) representation of *etr1* at the scale of 56 days; (d) first trapezoidal representation of the same process initially sampled with one point out of 30 (*etr30*).

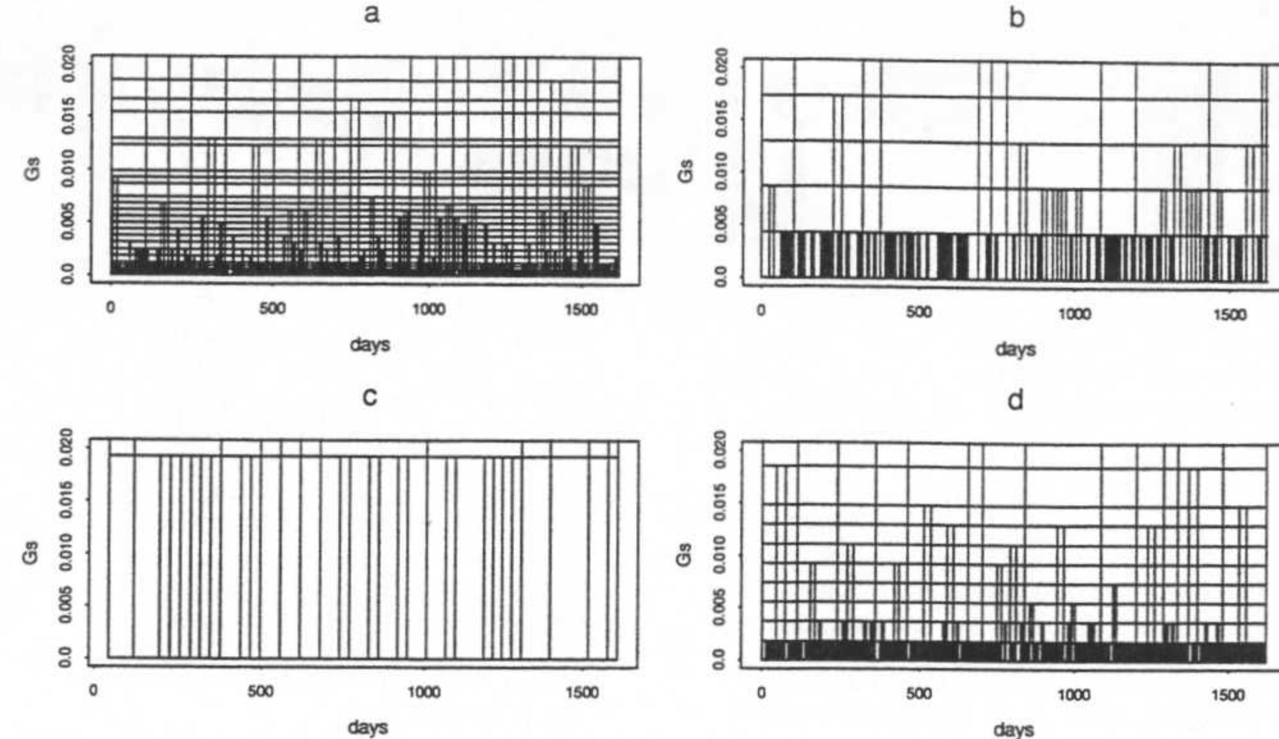

Fig. 8. Structures of granularities for *etr1* (a), *etr7* (b), *etr30* (c) and *etr3*(d).

The conclusion derived is that the process almost constantly oscillates. The measurement of consecutive durations and magnitude variations does not allow us to characterize those oscillations with more details. In such a process, oscillations may denote a random behaviour. To validate this hypothesis, the granularities spectrum expressed in number of trapezoidal episodes is plotted for the representation of *etr1* at the minimal scale of one day (see Fig. 9 b). A χ^2 test for a random distribution is successfully performed. This confirms the close relation between evapotranspiration and precipitation. It is worth noticing that even if the evolution lets 4 cycles clearly appear, a *temporal* random distribution is conceivable.

5.2. Embedded and skeleton evolutions for the prey-predator model with hunters

The original prey-predator model describes the evolution of prey and predator populations. It follows simple rules:

— when the predator population increases, the prey population decreases;
— when the predators are too numerous to feed correctly on the small number of remaining preys, their population decreases, which makes in turn the prey population increase;
— when the predators are few enough to feed on the existing number of preys, their population begins to increase, in turn the prey population begins to decrease.

We add to this classical feedback mechanism a population of hunters, assuming that they kill preys only, for 6 months a year, with a semi-sinusoidal shape of evolution. The simulation of the prey population is performed with a time-step of 10 days over a temporal domain of 30 years (see Fig. 10 a).

Intuitively, the fluctuations of the prey population are submitted to two influences: predators and hunters fluctuations. To verify this, let us examine the first trapezoidal representations of preys (see Fig. 10 b) and hunters.

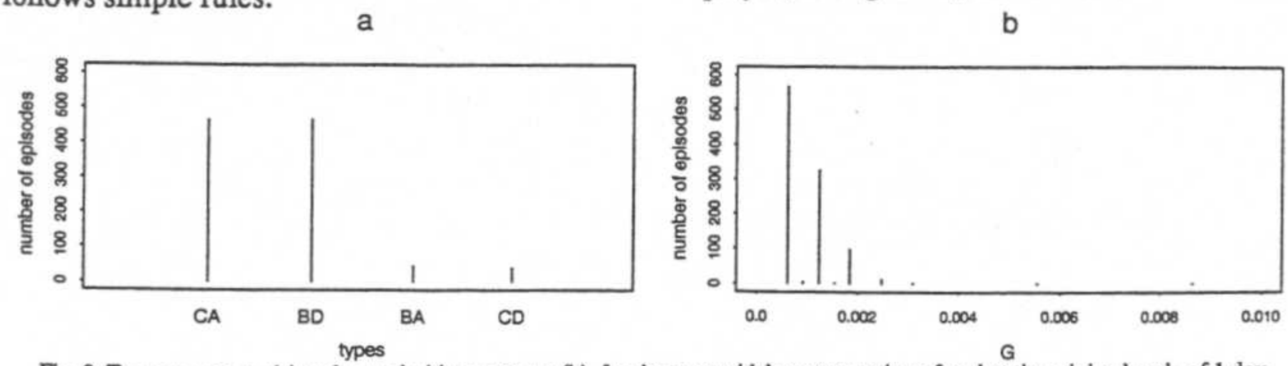

Fig. 9. Types spectrum (a) and granularities spectrum (b) for the trapezoidal representation of *etr1* at the minimal scale of 1 day.

The fluctuation indices over the whole domain are the same: there are approximately two episodes a year. Omitting hunters, the prey population evolution does not exhibit the 'fast' fluctuations represented by those two episodes. We conclude they are due to hunters' action. Abstractions successively performed eliminate them from the representations. In the structure of granularities (see Fig. 10 c), a large gap appears: the last representation with a non-abstracted episode is at the scale of 7 months; the next representation is at the scale of 28 years ! We thus extract the skeleton (see Fig 10 d) and embedded evolutions.

If we construct the trapezoidal representation of the skeleton evolution, its fluctuation index appears to match the predators evolution one. This verifies what we guessed. In the evolution of preys, 'fast' fluctuations come from the hunters fluctuations, and are embedded in 'slower' ones corresponding to the predators influence. We know precisely the time-scale beyond which highly-fluctuating influences can be disregarded for the representation (i.e. 7 months).

6. Discussion and conclusion

We built a system of interpretation of numerical evolutions obtained from simulation outputs (or observations curves). The idea of triangular or trapezoidal episode allows a description in terms of tendencies and curvatures followed by a qualitative and/or quantitative analysis. By simple geometrical constructions, trapezoidal episodes can be aggregated successively according to their duration. Duration is directly related to the concept of temporal granularity, thus the geometrical successive aggregation is a time-scale abstraction.

We chose to give a minor importance to numerical values than to qualitative shapes of evolution. However, for evaluating the time-interval over which a behaviour holds, we needed precise dating. Although temporal abstraction leads to a loss of behavioural precision (in terms of quantitative values, qualitative tendencies and curvatures), changes of granularity are perfectly located in the domain.

The case of 'embedded' evolutions is addressed, since it is frequently associated with time-scale problems. A technique of decomposition is proposed, that provides an help to distinguish several levels of fluctuations within the same evolution.

Several issues related to the abstraction procedure have to be pointed out:

— The boundary points of the whole evolution are kept in all the representations. In the most abstracted ones, these points can have an excessive influence on the global shape. This depends much on the nature of the process and can only be assessed by the user.

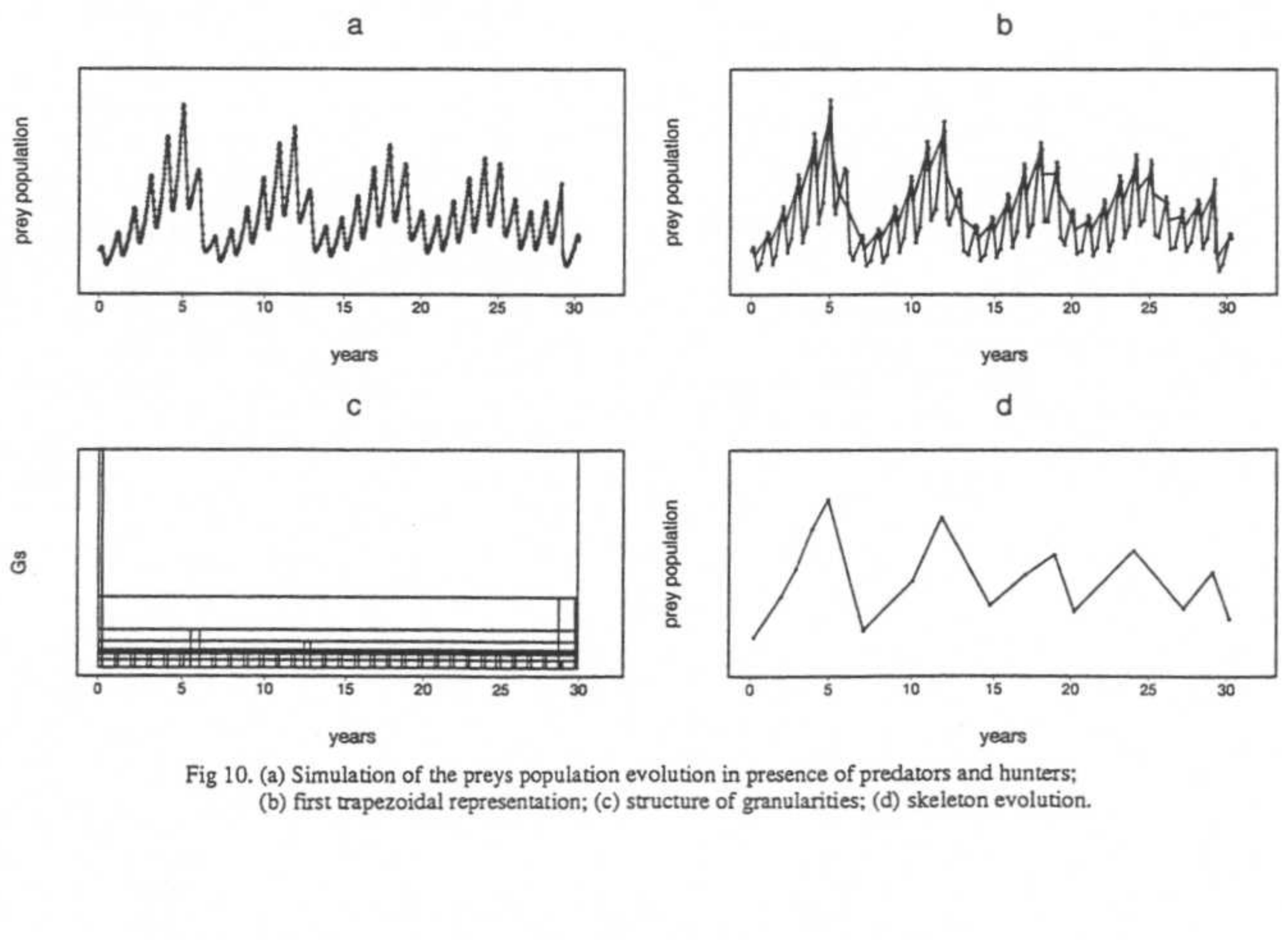

Fig 10. (a) Simulation of the preys population evolution in presence of predators and hunters; (b) first trapezoidal representation; (c) structure of granularities; (d) skeleton evolution.

— The fact that three episodes are needed to perform an abstraction is awkward when the episode of minimal duration is at the boundaries of the process. The currently implemented solution consists in looking for the next minimal episode, so the same remark as above applies with boundary episodes.
The sensitivity of representations to boundary points and episodes is assessed to be linked to the suitability of the initial sampling frequency: other initial series of points are tested and respective structures of granularities are compared (cf. § 4.2.).
Performing this comparison automatically and not by hand would necessitate the definition of various degrees of equivalence: number of episodes at the same minimal granularity, location of the minimal episodes, etc. Finally, comparing two distinct processes in terms of granularity appears feasible in a similar way.

Apart from the 'representation aspect' of the time-scale problem, it was required to help the choice among possible representations. For this task, user's needs (concerning the domain and the initial minimal granularity) are not the only things to be taken into account. The fluctuating features at each step of abstraction have to be analysed as completely as possible. Tools of interpretation were built in order to achieve a symbolic, numerical, and statistical description, and determine standard dynamic features. They are complementary to the simple visual facility of a geometrical description and allow one to assess the effect of abstraction over behavioural features.

A good perspective is to examine the ways used to compare two symbolic strings. Pattern-matching algorithms used for instance in biological sequences analysis should perfectly fit. As previously said for structures of granularities, this facility should be extended to compare two distinct processes at different levels of abstraction.

This analysis could be useful in design, as it may help the user to visualize how processes behave. It can also provide him with a guidance for changing the simulation time-steps, or modifying the model of a device. In this respect, coupling the current system to a knowledge-based system, incorporating both interpretation and decision rules, is a perspective.

Acknowledgements
This work is part of Laurent Ayrolles' Ph.D. research, jointly supported by INRA and the Framentec-Cognitech company (Paris). We thank particularly François Arlabosse (Framentec-Cognitech) for his interest in our work.

References

[1] HOBBS J. R.: 'Granularity'. Proc. AAAI, Los Angeles, CA, 1985

[2] STRUSS P.: 'On temporal abstraction in qualitative physics'. Preliminary Report, September 1993

[3] WELD D.: 'The use of aggregation in causal simulation', *Artificial Intelligence*, 1986, **30**, (1), pp. 1-34

[4] EUZENAT J.: 'Représentation granulaire du temps', *Revue d'intelligence artificielle*, 1993, **7**, (3), pp. 329-361

[5] MAÏM E.: 'Reasoning with different granularities in CRL'. Proc. IMACS, Dublin, Ireland, 1991

[6] KUIPERS B.: 'Abstraction by time scale in qualitative simulation'. Proc. AAAI, Seattle, WA, 1987

[7] IWASAKI, Y.: 'Reasoning with multiple abstraction models'. Proc. 4th Int. Workshop on Qualitative Physics, Lugano, Switzerland, July 1990

[8] RICKEL J., PORTER B.: 'Automated modeling for answering prediction questions: exploiting interaction paths'. Proc. 6th Int. Workshop on Q.R. about Physical Systems, Edinburgh, Scotland, 1992

[9] CHEUNG J. T.-Y., STEPHANOPOULOS G.: 'Representation of process trends -Part I. A formal representation framework', *Comp. Chem. Engng*, 1990, **14**, (4/5), pp. 495-510

[10] CHEUNG J. T.-Y., STEPHANOPOULOS G.: 'Representation of process trends -Part II. The problem of scale and qualitative scaling', *Comp. Chem. Engng*, 1990, **14**, (4/5), pp. 511-539

[11] AYROLLES L., GUERRIN F.: 'Towards a general approach to deal with temporal granularity'. Proc. IJCAI Workshop on Engineering Problems for Qualitative Reasoning, Chambéry, France, 1993

[12] WILLIAMS J.R., JONES C.A., DYKE P.T.: 'A modeling approach to determining the relationship between erosion and soil productivity', *Transactions of the ASAE*, 1984, **27**, pp. 129-144