At the proceedings of ESM'2006 The European Simulation and Modelling Conference, LAAS, Toulouse, France, October 23-26

Holistic Metrics, a Trial on Interpreting Complex Systems

J. Manuel Feliz-Teixeira
António E. S. Carvalho Brito

April 2006

GEIN – Section of Management and Industrial Engineering
Faculty of Engineering of the University of Porto, PORTUGAL
Emails: feliz@fe.up.pt; acbrito@fe.up.pt

KEYWORDS: Complex system, holistic metrics, time domain, frequency domain, base states, state projection.

ABSTRACT

In this article is proposed a simple method for estimating or characterize the behaviour of *complex systems*, in particular when these are being studied throughout simulation. Usual ways of treating the complex output data obtained from the activity (real or simulated) of such a kind of systems, which in many cases people classify and analyse along the *time domain*, usually the most complex perspective, is herein substituted by the idea of representing such data in the *frequency domain*, somehow like what is commonly done in *Fourier Analysis* and in *Quantum Mechanics*. This is expected to give the analyst a more holistic perspective on the system's behaviour, as well as letting him/her choose almost freely the *complex states* in which such behaviour is to be *projected*. We hope this will lead to simpler processes in characterizing complex systems.

1. Introduction

There are presently very few notes on the kind of metrics that could be reliable and of practical relevance when applied to the interpretation of complex systems behaviour. These systems are often based on intricate structures where a high number of entities interact with each other. Metrics are there for appropriately characterizing the nodes or individual parts of such structures, or small groups of them, but when the intent is a measure for the complete structure either they fail or appear to be too simplistic. That is certainly a good reason for modelling those cases using a *strategic* point of view, removing the *time* variable from the process, as in doing so the complexity is reduced *a priori*.

But when a dynamic and detailed representation is essential, the interpretation of the results and the characterization of the system frequently fail. This issue seems sometimes also related to a certain tendency impregnated in the minds to look at the systems from a pre-established perspective. At this point, however, perhaps this may be considered a conflict between different scientific approaches: the classical western reductionism, of anglo-saxonic inspiration, which believes the best approach is to break the system into small parts and understand, model or control those parts separately and then join them together, therefore looking at the world in an individualist way; and a more holistic approach, a vision slowly spreading and largely inspired by oriental cultures, which considers that each part of the system must be seen together with the whole and not in isolation, and therefore locates the tone in how the interactions between such parts contribute to the whole behaviour. Hopp & Spearman (2001, pp.16), for instance, comment about this saying that "*too much emphasis on individual components can lead to a loss of perspective for the overall system*".

A significant number of authors defend this opinion, pointing out the importance of developing a more holistic point of view to interpret and study systems behaviour, in a way that analyses maintain enough fidelity to the system as a whole. As

Tranouez et al. (2003), who apply simulation to ecosystems, would say: a complex system is more than the simple collection of its elements.

In management science, for instance, the "western" approach frequently generates difficulties at the *interfaces* between elements, typically of inventory or communication type. On the other hand, as *just-in-time* (JIT) systems give better emphasis to the relations and interactions and are continuously improving, the overall movements tend to be more harmonious. JIT already looks at systems in a certain *holistic* way. The same seems to be true in regard to other fields where simulation is applied, and mainly when the number of states to simulate is high.

2. Holist measuring (a proposal)

But, what concerning metrics? How can one measure such a high number of states typically found in complex systems in order to effectively retrieve from them some sort of useful information?

As a metric is a *characterization*, we could think that maybe the modern *Data Mining* (DM) techniques could be extensively applied, for instance. These techniques use decision trees and other algorithms to discover hidden patterns in huge amounts of data, and are nowadays applied to almost any problem based on extensive data records, for instance, in *e-Commerce* for customer profile monitoring, in genetics research, in fraud detection, credit risk analysis, etc., and even for suspected "terrorist" detection (see Edelstein, 2001; Edelstein, 2003). However, they often imply the usage of high performance computers, sometimes with parallel processors, as well as huge computational resources to analyse *GBytes* or even *TBytes* of data. They are useful when any single record of data can be precious for the future result, and thus when all data *must* be analysed.

On the other hand, in many practical simulations a significant amount of data is not significant for the final conclusions, the simulation process is in itself a filter, and therefore such data may well be ignored in the outputs, even if it could have been essential to ensure the detailed simulation process to run. In the perspective of the author, maybe there is a way that could deserve some attention: the idea is to filter such data during the simulation execution and, at the same time, to turn the measures probabilistic by using an approach somehow inspired by the *Fourier Analysis* and the *Quantum Mechanics*. That is, to represent the overall *system state* (ψ) in terms of certain *base functions* (ψ_i), and then to measure the *probabilities* (α_i) associated with each of these functions. The interesting aspect of this is that each base state function (ψ_i) could even be arbitrarily chosen by the analyst, and the probabilities (α_i) easily computed during the simulation process. Final results would then be summarised in some expression of the form:

$$\psi = \alpha_1 \psi_1 + \alpha_2 \psi_2 + \dots \alpha_j \psi_j + \dots \alpha_n \psi_n \qquad (1)$$

which could be interpreted as: there is a probability of α_1 that the system will be found in the state ψ_1, a probability of α_2 that the system will be found in the state ψ_2, etc. This would be the final measure of the system, in a sort of characterization of expectations under certain conditions. This also corresponds to *projecting* the system behaviour into the generalised vectors base of *state functions* (ψ_i). The amounts α_i simply correspond to the values of those projections.

In *Fourier Analysis*, for instance, the complex behaviour observed in the *time axis* (see the example of figure 1) is substituted by the decomposition of such a signal into *sine* and *cosine* mathematical functions, and that way transferred to the *frequency* domain.

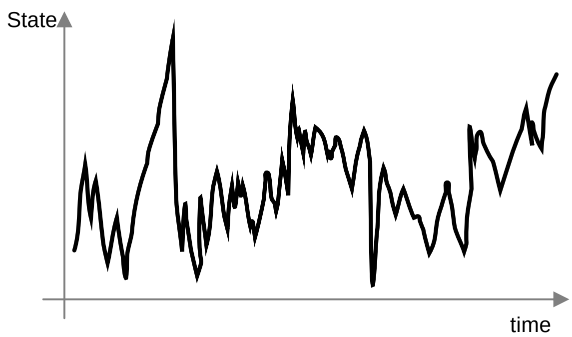

Fig. 1 Example of a general complex signal

The result is that the analyst is now much more able to visualize and to interpret the complexity of the previous signal, since it is as if this signal would be now expressed in terms of *patterns* (see example of figure 2). What firstly appeared as a confusing and almost randomly up-and-down behaviour may now be simply understood as the summation of some sinusoidal patterns with different amplitudes. *Quantum Mechanics* uses a similar formalism. We believe that the method proposed here will help generating such a clean view also when applied to the behaviour of *complex systems*.

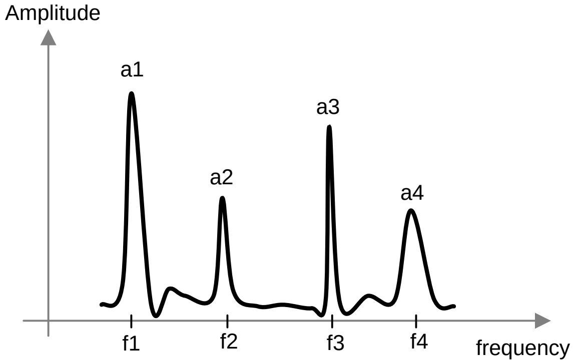

Fig. 2 Typical signal in the frequency domain

The present proposal may also be understood as an attempt to represent the system behaviour in terms of a sort of generalised histogram, where the *categories* are the functions ψ_i, which may correspond to the frequencies f_i in the previous figure, and the probabilities α_j are made to correspond to the amplitudes a_j in the same figure. In terms of this figure, the analyst would recognize a probability of a_1 that the system would be found in the state f_1, a probability of a_2 that the system would be found in the state f_2, etc.

3. An imaginary example

To help explain this, we can imagine a complex system like the Supply Chain shown in figure 3, for example.

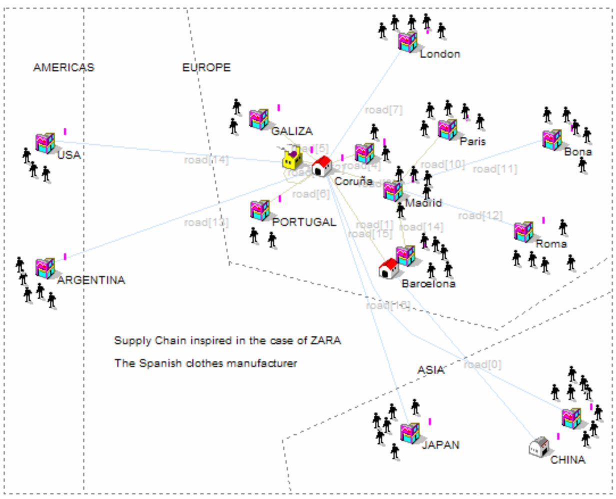

Fig. 3 Imaginary Supply Chain inspired by ZARA

This is an example inspired by the company ZARA, the trendy Spanish clothes manufacturer of La Coruña. This company, from the INDITEX group, is worldwide known as a paradigm of success, despite its owner, and major manager, Mr Ortega, the second richest person in Spain, refusing several conventional practices claimed by most schools of management. ZARA refuses, for instance, the idea of advertisement. Forgive me if indirectly I am advertising it here.

Returning to our subject, how can we apply our concept of holistic metrics to retrieve some useful information from such a complex case[1]? How can we specify the *base functions* (or *base states*) in which the system's behaviour will be *projected*? How will we calculate and represent the respective projections?

First of all, we have to choose the ψ_i functions into which the measures will be *projected*. We may choose them in terms of some specific *conditions* related to the information that must be obtained from the system. For example, if Mr Ortega is concerned about the levels of *stockouts*, *holding costs*, *service level*, *turnover*, etc., which are typical measures of Supply Chain Management, he may for example define some sort of base functions by using conditions of the type:

ψ_1 – *Stockouts above 7%*
ψ_2 – *Holding costs above 5%*
ψ_3 – *Service level under 75%*
ψ_4 – *Turnover under 2*

Then, while the system is running, it must be *projected* into these set of functions, that is, the occurrences of each of these conditions must be counted up, whenever they are true.

Supposing n_j the accumulated number of occurrences of the condition ψ_j, and N_j the total number of its samples, an estimation of α_j can simply be computed as:

$$\alpha_j = n_j / N_j \tag{2}$$

And the overall system state will therefore be expressed as:

$$\psi = (n_1/N_1)\psi_1 + (n_2/N_2)\psi_2 + (n_3/N_3)\psi_3 + (n_4/N_4)\psi_4 \tag{3}$$

Notice that, in general, *base functions* are

[1] In this figure is represented less than perhaps 10% of the real ZARA global Supply Chain structure.

chosen to be orthogonal, or independent of each other, but in fact that is not a *must* for using this type of representation. One can also *project* a system into *non orthogonal* axis. As we said previously, such a measure may be seen as a characterization of expectations under certain conditions. The overall system state is, in reality, represented by the following weighted equation:

α_1 *(Stockouts above 7%)* + α_2 *(Holding costs above 5%)* + α_3 *(Service level under 75%)* + α_4 *(Turnover under 2)* (4)

Now, if we build a histogram out of this data, we will *characterize* the system by means of a probabilistic graphical format, obtaining something of the type:

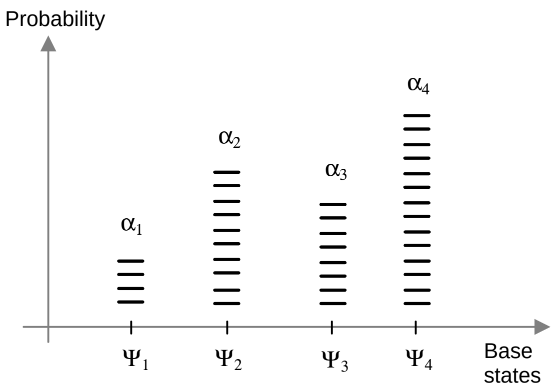

Fig. 4 Characterization of the system's behaviour

Were the probabilities are the α_i.

So, once the *base states* are well defined by the analyst, the characterization of the system is possible, no matters how complex the system is. We recall that in many practical cases the analyst is mainly focused in being sure that certain variables of the model do not cross some upper or lower limits, or, if they do, with which probability it happens.

In order to evaluate the system in a wider range of modes of behaviour, several studies of this kind can be made with the system operating in different conditions. That will make possible to improve the knowledge about the system, or its characterization.

The former example was taken from a typical Supply Chain problem (see Feliz-Teixeira, 2006, pp. 222), but this technique can be applied in general to other complex systems. For example, in a traffic system of a town, the *complex states* could be chosen to be the number of cars exceeding a certain value in a certain region, the travel time exceeding a certain value in another region, the number of public vehicles reaching a certain zone inferior to the minimum required, etc. As we recommend that these base functions (or *complex base states*) be well defined before simulation takes place, it implies that the simulation objectives must be well known prior to the start of the simulation process. Not always this is possible, of course, since simulation can be used to detect anomalous situations not predictable by means of other methods, for example.

This technique may, however, be also used as a method for analyse any sort of results, by being directly applied to the raw outputs of the complex system. In that case, the simulation will be a standard process and all the work is done by data manipulation. The results, in principle, will be the same, but that approach will in general be much more time consuming.

Finally, we would like to emphasise that we use the term "*holistic metric*" for distinguishing this kind of approach from those approaches which usually characterize systems by means of *averages* and *standard deviations* taken over a certain number of variables (usually a high number). These, as we know, frequently confuse the analyst's mind with the complexity of the results, instead of allowing a useful interpretation of the system's behaviour. Quantity of information is not all, and sometimes it can even generate confusion instead of clarity, if it is in excess. Besides, the method presented here goes on the trend of the "holistic" mind that seems to emerge in our days, as we defend.

4. CONCLUSIONS

Complex results generated by a *complex system* are very much dependent on how the analyst looks at the system and on how such results are analysed. We would say that any complex system can be minimally understood as long as the analyst knows what to search for, that is, if the objectives of the study are previously defined. This is because such objectives can in reality be used to establish the *base functions* (vectors) of an imaginary space where the complex behaviour will be *projected*, that way giving an automatic meaning to the results.

This may also be seen as an attempt to measure the outputs of systems in the *frequency domain* (as in *Fourier Analysis* and in *Quantum Mechanics*), instead of in the *time domain* where signals usually are more difficult to interpret. Although no practical cases have yet been studied based on the idea presented in this article, we expect to use and test this approach in our next studies of simulation. We would also be pleased with receiving some feedback from anyone who decided to apply the same logic.

References:

Edelstein, H. (2001). Pan For Gold In The Clickstream. *InformationWeek.com*.

Edelstein, H. (2003). Using Data Mining to Find Terrorists. *Data Mining Review*.

Feliz-Teixeira, J. M. (2006). *Flexible Supply Chain Simulation (thesis).* Faculty of Engineering of Universitry of Porto, Portugal.

Hopp, W. J., & Spearman, M. L. (2001). *Factory Physics* (second ed.): Irwin McGraw-Hill.

Tranouez, P., Lerebourg, S., et al. (2003). *Changing the Level of Description in Ecosystem Models: an Overview.* Paper presented at the The 2003 European Simulation and Modelling Conference, Naples, Italy.