Discovering event prediction knowledge

Marlon Núñez, Rafael Morales and Raúl Fidalgo

Departamento de Lenguajes y Ciencias de la Computación, Universidad de Málaga,
Campus de Teatinos s/n, Málaga 29071, Spain
{mnunez,morales,fidalgo} @lcc.uma.es

Abstract. Discovering temporal patterns from event sequences is a difficult task. Many problems of the real life may have different behaviours ranging from regular to chaotic (i.e. weather, stock market). By definition, the chaotic phenomena cannot be predicted in the long-term, nevertheless short-term prediction is possible. A recent study constructs learning examples from incoming events irregularly distributed over time, grows special regression trees with temporal constraints, and composes a prediction model of the observed system. This paper investigates this approach in a greater depth and makes several improvements. It introduces a novel technique to detect and handle chaotic and nonchaotic behaviours. It also introduces automatic discovery of the observation windows, called Latency Windows. These new functions have been successfully validated with systems whose behaviours (chaotic and nonchaotic) are known using differential equations.

1 Introduction

If we know the differential equations of a dynamical system we may predict its behaviour. But if the system is nonlinear with chaotic behaviour this prediction is impossible for the long-term, however a short-term prediction is possible. A learning system should detect chaos in order to know the validity of the discovered patterns.

An important problem is to detect chaos in a system. If differential equations are known, the problem may be faced using several known methods from chaos theory (Lyapunov exponents, fractal dimensions, among others) [3], [12]. In most real world situations we do not know the differential equations and so we must calculate the exponents from univariate time series.

This paper summarizes the BPL method [7], [9] for learning event prediction knowledge. This method constructs learning examples from incoming events, grows special regression trees with temporal constraints, and composes a model of the observed system to predict the time of occurrence of future target events. This paper makes several improvements to this approach. It introduces a novel technique to detect and handle both chaotic and nonchaotic behaviours in a multivariate way. It also introduces automatic discovery of the observation windows, called Latency Windows. Chaos theory establishes that temporal patterns found from chaotic intervals are valid only during that chaotic interval and for a short period of time. This paper describes how BPL uses this principle.

Previous researches related to this field are the discovery of patterns of sequences [4], [11] as well as time series analysis [13], [3]. Sequence discovery methods searches for frequent *episodes* or *subsequences* of events. Sequences do not handle chaotic behaviours. In time series analysis, the only way to detect and deal with chaos is using univariate approaches. Another important difference with current approaches is that BPL uses static information of the observed system and its environment to build the temporal patterns.

In this paper, we applied the new facilities of BPL for predicting short- and mid-term events in chaotic systems. First of all, in section 2.1, we summarise how BPL method handles chaotic and nonchaotic behaviours. In section 2.2 we explain in some detail the basic concepts used in the paper. Section 2.3 describes the algorithms. In section 3, we present several experiments with chaotic systems. Section 4 presents the conclusions.

2 BPL method

In order to analyse the possible consequences of an event, the following natural assumptions are made: an event could be a consequence of some temporal correlation of past events, static characteristics of the observed system, and static characteristics of the environment; on the other hand, an event could have an effect in the future for a limited period of time. Because of the above, two observed systems in the same state situated in two different environment conditions behave differently. For this reason BPL also learns from events and static information of the system and its environment.

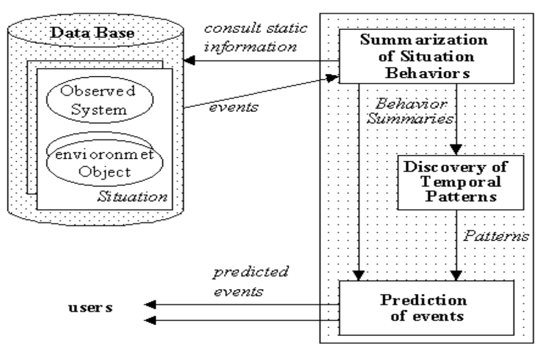

Fig. 1. Behavior Pattern Learning (BPL) processes.

2.1 Overview

Figure 1 shows the three processes of BPL. The first process is the *summarisation of situation behaviours*. A situation is made up of an *observed system* and its *environment*. Both are characterised by dynamic attributes and static attributes. A change in the value of a dynamic attribute, as well as the time in which this change took place, is registered as an event. An event schedules the construction of several behavior summaries in the future during its latency window.

Figure 2 illustrates the way BPL monitors the consequences of *eventX*, scheduling four BehaviorSummaries during a time interval. This time interval is the period of time during which *eventX* is supposed to affect the future. BPL constructs summaries with static and dynamic attributes of the system and its environment, as well as new features from events, such as the duration of current event values, repetitions of an event in a period of time, and oldness of past events.

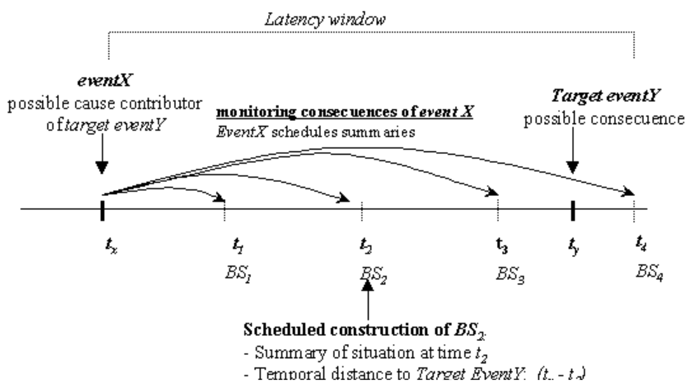

Fig. 2. Monitoring the consequences of an event

The second process is the *discovery of temporal patterns*. Chaotic systems are characterised by continuous increment of the entropy. BPL search for these intervals applying a novel multivariate technique that uses the discovered patterns as elements of order/disorder. If the prediction is to be made in a chaotic interval, then temporal patterns are constructed taken into account only the behaviours summaries in that chaotic interval. If the prediction is to be made in a nonchaotic interval, then the temporal patterns are constructed taking into account the whole history of nonchaotic behaviours summaries. Temporal patterns show: the preconditions (static and dynamic) for a target event to happen, the next expected time interval of occurrence and the support of the pattern. This process constructs a set of temporal patterns for predicting a target event. This set of patterns is called a behaviour tree (see Figure 3).

The third process is called the *prediction of events*. This process receives a behaviour summary (without class) and drop it down the behaviour tree to predict an event.

Figure 3 illustrates a small part of a behaviour tree for predicting the most probable interval during which an HTTP service is expected to fail in a Server. Figure 3 shows that IF the operating system is Linux AND the duration of the available memory in the band "5%-12%" is from 1 to 10 minutes, and the number of times the CPU utilisation has reached its maximum “90%-100%” is from 1 to 4 in the last 5 minutes, then a Failure "Service Degradation=TypeA" is expected in 8 to 25 minutes. The support of this prediction is 0.4%. If the number of times the CPU has reached its maximum is from 5 to 16 in the last 5 minutes, then a “Service Degradation=TypeA" is expected in 2 to 9 minutes.

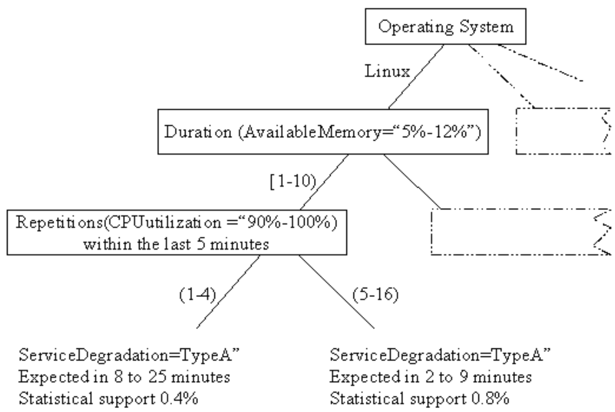

Fig. 3. A behaviour tree

3 Experimentation

In this section we describe the experiments for validating and applying the proposed method and present the results. In section 3.1 and 3.2 we empirically validated BPL model for predicting target events of variables with chaotic and nonchaotic behaviour.

We did experiments with the nonlinear chaotic system: the Double Well Oscillator. In the system, there is a mass *m* that vibrates on a filament. The mass is exerted to an electromagnetic field. In presence of certain conditions the period of the vibration increments, becoming chaotic. Figure 4 shows the system where *m* is the mass and *x(t)* is the displacement of the mass.

The objective is these experiments is to observe the behaviour of the system during a period of time in presence of different conditions in order to discover a model of dynamic behaviour using behaviour trees and predict the occurrence of certain positions of the object The double well oscillator can be modelled mathematically by a differential equation, called Duffing equation. If its parameters are configured appropriately, including a force $F(t) = F_o \cos wt$ exerted to the mass,

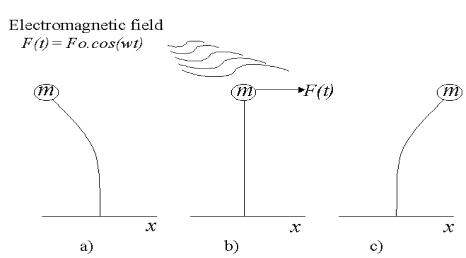

Fig. 4. The double well oscillator. *x* is the displacement of the mass *m.*

we have a differential equation that shows several behaviours [6]: $x''+x'-x+x^3 = F_o cos(t)$. For the double well oscillator we discretized $x(t)$ and $x'(t)$. We use 10 thresholds for the variables and 3 thresholds for the derivatives.

For measuring the error we used the Relative Mean Square Error (RMSE), a typical measure used in regression. It is relative because it is the ratio between the error of the analysed predictor and the error of a naive media predictor. Error of a predictor is measured calculating the sum of square distances between the analysed predictor and the real value. If RMSE is greater than 1, it means that the analysed predictor is worst than the naive media predictor. Good RMSErrors obtained by regression algorithms for typical problems fall bellow 0.3.

In experiment 1 we learn and test patterns within a limited chaotic interval *[0..5000 seconds],* applying 10-crossvalidation. In the experiment 2 we take the resulting behavior tree of experiment 1 for predicting target events in the future in the chaotic problem.

3.1 Experiment 1: Learning and testing patterns within a limited chaotic interval

The purpose of the experiment 1 is to discover short and long-term patterns within a limited chaotic time interval. We will use 10-crossvalidation for measuring RMSError. In experiment 2 we will take the long-term prediction behavior tree (the one with closest error to global cross-validation) for measuring prediction error in the future.

With that purpose in mind, we made BPL find patterns in a chaotic situation with two dynamical attributes $x(t)$ and $x'(t)$. $x(t)$ is the solution of differential equation (1) with F_0=0.8. We selected a frequent event type as short-term target event. This event type was x=A. We selected x=e, a rare event type, as long-term target event. Table 1 also shows the effect of Latency Windows by calibrating the parameter *Lwf (Latency Windows extension factor)* as explained to section 3.5. This table shows low prediction errors. The best results were obtained with *lwf =3*.

Table 1. Error estimation for short/long term prediction in chaotic situation x(t), x'(t) for F_0=0.8 within limited interval

	Lwf (Latency Windows extension factor)	Number of Behavior Summaries	Number of leaves of the tree with error nearest to crossvalidated	10-Cross-validated RMSError
Short-term target event	2	11231	446	0,120
	3	11157	459	0,056
	4	11059	478	0.0729
Long-term target event	2	6192	104	0,044
	3	6012	102	0,030
	4	6002	98	0,041

3.2. Experiment 2: testing long-term prediction patterns in the future.

In experiment 2 we take the resulting behavior tree of experiment 1 for predicting long-term target events in the future of the same chaotic situation. Table 2 shows that error increases dramatically from 0.08, obtained immediately after experiment 1 learning interval, to 2.78, obtained at the interval [14000..14500] seconds. This experiment shows that long-term prediction of chaotic systems not possible as it is stablished in chaos theory (Hilborn R., 2000) (Stewart I., 1997). Remember the if RMSE is greater than 1, it means that the analysed predictor is worst than the naive media predictor.

Table 2. Testing patterns of experiment 1 in the *far* future.

Testing in the chaotic interval:	RMSError
[5000..5500] seconds	0.08
[8000..8500] seconds	0.79
[11000..11500] seconds	1.92
[14000..14500] seconds	2.78

4. Conclusions

BPL may be used as a tool for finding general temporal patterns that are consistent with past experiences, which may allow us to understand the phenomena in the present, and try to predict the future. It is more similar to the way experts express their knowledge

Validation of BPL was done by learning and testing behavior trees in limited (closed) time interval. Experiment 1 and 2 showed that:

- BPL found patterns for predicting short and mid-term target events with RMSErrors bellow 0.10,
- Patterns obtained in a closed chaotic interval are not valid outside this interval.
- Generated knowledge that may be reviewed by users.
- Automatic calculation of latency windows proved to be helpful. Best result was obtained with factor *Lwf* of 3.

Acknowledgments

We would like to thank Prof. Ivan Bratko who suggested us the use of BPL with differential equations, which became the best way to validate our method. We also thank Ignacio Ballesteros for developing some modules. This work has been partially supported by project FACA PB98-0937-C04-01 of the CICYT, Spain. FACA is part of the FRESCO project

References

1. Atmanspacher, H., Scheingraber, H., (1987). A fundamental link between system theory and statistical mechanics, Foundations in Physics, 17, 939-969.

2. Breiman L., Friedman J.H., Olsen, R.A., & Stone, C.J., (1984). *Classification and Regression Trees*, Wadsworth Int. Group, Belmont, California.
3. Hilborn R. (2000) Chaos and Nonlinear Dynamics. Oxford University Press, 2000.
4. Manila, H., Toivonen H., and Verkamo Inkeri (1997) Technical Report C-1997-15, Department of Computer Science, University of Helsinki.
5. Mora LL., Núñez M., Morales R., Triguero F., in Survey on Event Prediction Methods for Time Series, submitted to *ECML-2001*.
6. Moon, F. C., & Holmes, P. J. (1979). A magnetoelsastic strange attractor. *Journal of Sound Vibration*, 65, 285
7. Núñez, M., Morales R., Triguero F. (2002). Automatic Discovery of Rules for Predicting Network Management Events. *IEEE Journal on Selected Areas in Communications*, vol. 20, 4, 2003.
8. Núñez, M. (2000). Generalised Regression Trees, *14th Intl. Conference of Statistical Computing (Compstat), Physica-Verlag, 367-372,* The Netherlands.
9. Núñez, M. (2000). Learning Patterns of Behaviour by Observing System Events, in *Lecture Notes in Artificial Intelligence, ECML 2000,* Springer-Verlag.
10. Philips, P. (1998), "Impulse Response and Forecast Error Variance Asymptotic in Nonstationary VARs", Journal of Econometrics 83, pp 21-56.
11. Srikant, R. and Agrawal R (1996). Mining Sequential Patterns: Generalisation and Performance Improvements. In *Proceedings -96*, Avignon, France.
12. Stewart I., 1997. Does God Play Dice? The New Mathematical of Chaos, Penguin.
13. Box, G.E.P., Jenkins, G.M. Time Series Analysis forecasting and control. Prentice Hall, USA, 1976.