Artificial Intelligence 262 (2018) 163–188

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

On the adoption of abductive reasoning for time series interpretation

T. Teijeiro *, P. Félix

Centro Singular de Investigación en Tecnoloxías da Información (CITIUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

A R T I C L E   I N F O

Article history:
Received 31 March 2017
Received in revised form 10 November 2017
Accepted 4 June 2018
Available online xxxx

Keywords:
Abduction
Interpretation
Time series
Temporal abstraction
Temporal reasoning
Non-monotonic reasoning
Signal abstraction

A B S T R A C T

Time series interpretation aims to provide an explanation of what is observed in terms of its underlying processes. The present work is based on the assumption that the common classification-based approaches to time series interpretation suffer from a set of inherent weaknesses, whose ultimate cause lies in the monotonic nature of the deductive reasoning paradigm. In this document we propose a new approach to this problem, based on the initial hypothesis that abductive reasoning properly accounts for the human ability to identify and characterize the patterns appearing in a time series. The result of this interpretation is a set of conjectures in the form of observations, organized into an abstraction hierarchy and explaining what has been observed. A knowledge-based framework and a set of algorithms for the interpretation task are provided, implementing a hypothesize-and-test cycle guided by an attentional mechanism. As a representative application domain, interpretation of the electrocardiogram allows us to highlight the strengths of the proposed approach in comparison with traditional classification-based approaches.

1. Introduction

The interpretation and understanding of the behavior of a complex system involves the deployment of a cognitive apparatus aimed at guessing the processes and mechanisms underlying what is observed. The human ability to recognize patterns plays a paramount role as an instrument for highlighting evidence which should require an explanation, by matching information from observations with information retrieved from memory. Classification naturally arises as a pattern recognition task, defined as the assignment of observations to categories.

Let us first state precisely at this point what is the problem under consideration: we wish to interpret the behavior of a complex system by measuring a physical quantity along time. This quantity is represented as a time series.

The Artificial Intelligence community has devoted a great deal of effort on different paradigms, strategies, methodologies and techniques for time series classification. Nonetheless, in spite of the wide range of proposals for building classifiers, either by eliciting domain knowledge or by induction from a set of observations, the resulting classifiers behaves as deductive system. The present work is premised on the assumption that some of the important weaknesses of this approach lie in its deductive nature, and that an abductive approach can address these shortcomings, which are described below.

* Corresponding author.
E-mail address: tomas.teijeiro@usc.es (T. Teijeiro).

https://doi.org/10.1016/j.artint.2018.06.005

Let us remember that a deduction contains in its conclusions information that is already implicitly contained in the premises, and thus it is truth-preserving. In this sense, a classifier ultimately assigns a label or a set of labels to observations. This label can designate a process or a mechanism of the system being observed, but it is nothing more than a term that summarizes the premises implied by the observations. Conversely, abduction, or inference to the best explanation, is a form of inference that goes from data to a hypothesis that best explains or accounts for the data [21]. Abductive conclusions contain new information not contained in the premises, and are capable of predicting new evidence, although they are fallible. Abductions are thus truth-widening, and they can make the leap from the language of observations to the language of the underlying processes and mechanisms, responding to the aforementioned problem in a natural way [24]. For example, consider a simple rule stating that if a patient experiences a sudden tachycardia and a decrease in blood pressure, then we can conclude that he or she is suffering from shock due to a loss of blood volume. From a deductive perspective, *loss of blood volume* is just a name provided by the rule for the satisfaction of the two premises. However, from an abductive perspective, *loss of blood volume* is an explanatory hypothesis, a conjecture, that expands the truth contained in the premises, enabling the observer to predict additional consequences such as, for example, pallid skin, faintness, dizziness or thirst.

Of course, the result of a classifier can be considered as a conjecture, but always from an external agent, since a classifier is monotonic as a logical system and its conclusions cannot be refuted from within. Classifier ensembles aim to overcome the errors of individual classifiers by combining different classification instances to obtain a better result; thus, a classifier can be amended by others in the final result of the ensemble. However, even an ensemble represents a bottom-up mapping, and classification invariably fails above a certain level of distortion within the data. The interpretation and understanding of a complex system usually unfolds along a set of abstraction layers, where at each layer the temporal granularity of the representation is reduced from below. A classification strategy provides an interpretation as the result of connecting a set of classifiers along the abstraction structure, and the monotonicity of deduction entails a propagation of errors from the first abstraction layers upwards, narrowing the capability of making a proper interpretation as new abstraction layers are successively added. Following an abductive process instead, an observation is conjectured at each abstraction layer as the best explanatory hypothesis for the data from the layer or layers below, within the context of information from above, and the non-monotonicity of abduction supports the retraction of any observation at any abstraction layer in the search for the best global explanation. Thus, bottom-up and top-down processing complement one another and provide a joint result. As a consequence, abduction can guess the underlying processes from corrupted data or even in the temporary absence of data.

On the other hand, a classifier is based on the assumption that the underlying processes or mechanisms are mutually exclusive. Superpositions of two or more processes are excluded; they must be represented by a new process, corresponding to a new category which is different and usually unrelated to previous ones. Therefore, an artificial casuistry-based heuristics is adopted, increasing the complexity of the interpretation and reducing its adaptability to the variability of observations. In contrast, abduction can reach a conclusion from the availability of partial evidence, refining the result by the incremental addition of new information. This makes it possible to discern different processes just from certain distinguishable features, and at the end to infer a set of explanations as far as the available evidence does not allow us to identify the best one, and they are not incompatible with each other.

In a classifier, the truth of the conclusion follows from the truth of all the premises, and missing data usually demand an imputation strategy that results in a conjecture: a sort of abducing to go on deducing. In contrast, an abductive interpretation is posed as a hypothesize-and-test cycle, in which missing data are naturally managed, since a hypothesis can be evoked by every single piece of evidence in isolation and these can be incrementally added to reasoning. This fundamental property of abduction is well suited to the time-varying requirements of the interpretation of time series, where future data can compel changes to previous conclusions, and the interpretation task may be requested to provide the current result as the best explanation at any given time.

Abduction has primarily been proposed for diagnostic tasks [10,33], but also for question answering [15], language understanding [22], story comprehension [6], image understanding [36] or plan recognition [28], amongst others. Some studies have proposed that perception might rely on some form of abduction. Even though abductive reasoning has been proven to be NP-complete or worse, a compiled form of abduction based on a set of pre-stored hypotheses could narrow the generation of hypotheses [24]. The present work takes this assumption as a starting point and proposes a model-based abductive framework for time series interpretation supported on a set of temporal abstraction patterns. An abstraction pattern represents a set of constraints that must be satisfied by some evidence in order to be interpreted as the hypothetical observation of a certain process, together with an observation procedure providing a set of measurements for the features of the conjectured observation. A set of algorithms is devised in order to achieve the best explanation through a process of successive abstraction from raw data, by means of a hypothesize-and-test strategy.

Some previous proposals have adopted a non-monotonic schema for time series interpretation. TrenDx system detects significant trends in time series by matching data to predefined trend patterns [19,20]. One of these patterns plays the role of the expected or normal pattern, and the other patterns are fault patterns. A matching score of each pattern is based on the error between the pattern and the data. Multiple trend patterns can be maintained as competing hypotheses according to their matching score; as additional data arrive some of the patterns can be discarded and new patterns can be triggered. This proposal has been applied to diagnose pediatric growth trends. A similar proposal can be found in [27], taking a step further by providing complex temporal abstractions, the result of finding out specific temporal relationships between a set of significant trends. This proposal has been applied to the infectious surveillance of heart transplanted patients. Another example is the Résumé system, a knowledge-based temporal abstraction framework [42,39]. Its goal is to provide