A First-Order Interval Temporal Logic for Adjacent Variables Temporal Data

Guido Sciavicco[1]

[1] University of Ferrara, Italy

Abstract

Multivariate time series are a very common non-tabular type of data. In many practical cases, multivariate time series encode real-world situations that include temporal information, and, recently, machine learning from datasets of multivariate time series has become a very active area of research. Modal symbolic learning has shown itself to be a serious alternative to sub-symbolic methods such as neural networks for non-tabular data; when applied to multivariate time series, modal symbolic learning makes use propositional temporal logic such as interval temporal logic. In special cases, however, multivariate time series display an internal structure that propositional modal logics are unable to capture. In this paper, we propose a first-order extension of propositional interval temporal logic, we describe its syntax and semantics, and we study its expressive power in relationship with such special cases of multivariate time series.

Keywords

interval temporal logic, expressive power, machine learning

1. Introduction

Let $\mathbb{D} = \langle\{1, 2, \ldots, N\}, <_t, =_t\rangle$ be a *finite linear order* (the *domain*) of *size* N, and let $V : \mathbb{D} \to \mathbb{R}$ be a *temporal variable* or *time series*; also, let $\mathcal{V}$ be a vocabulary of temporal variables. A *multivariate* time series T is a collection $T = \{V_1, \ldots, V_n \mid V_i \in \mathcal{V}, 1 \leq i \leq n\}$ of time series. A *temporal dataset* $\mathcal{T}$ is a collection $\mathcal{T} = \{T_1, \ldots, T_m\}$ of multivariate time series. Multivariate time series occur naturally in several real-world situations, from industrial data of devices and machines whose behaviour is monitored via sensors of various kinds, to position data of moving objects, to medical data of patients under observation, among many others.

In the machine learning realm, one important distinction is that among *symbolic* and *sub-symbolic* learning. The latter, in particular, encompasses the plethora of algorithms and methods that learn from datasets using functional representation of knowledge, from simple regressions, to support vector machines, to neural networks in different flavours. Symbolic learning, on the contrary, is characterized by learning explicit knowledge in logical form, and models for symbolic learning range from decision trees to set of rules, in several variants [1, 2, 3, 4, 5, 6, 7]. Since symbolic methods are logic-based, the logic that is used in a particular learning exercise can be seen as a parameter of the method. Until very recently, symbolic learning has been

OVERLAY 2023: 5th Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, November 7, 2023, Rome, Italy
guido.sciavicco@unife.it (G. Sciavicco)
0000-0002-9221-879X (G. Sciavicco)

essentially based on *propositional logic*, with very few exceptions, and consequently limited to tabular data. In the past few years, however, *modal* symbolic learning has been proposed as a generalization of symbolic learning to *modal logic*, and applied to non-tabular data. Modal logics have the ability of capturing a greater fraction of the internal structure of the instances of a non-tabular dataset, and by enriching intelligent symbolic models, such as decision trees, with the possibility of learning modal logic rules, one is able to extract such structure, as well as to learn and express interesting and complex patterns. In the temporal case, for example, modal symbolic learning is instantiated to *temporal* symbolic learning, and learning from temporal datasets is accomplished with, for instance, *temporal decision trees* [8], which are able to extract temporal patterns and express them using a suitable modal temporal logic. Among all possible modal temporal logics, *interval temporal logic* (of which several versions exist) revealed itself to be a a very useful tool in this regard. In a nutshell, one first sets up a set of *feature extraction functions* (e.g., the *mean* of a set of real numbers); then one builds a set of propositional letters, each resulting from the comparison between the result of applying one feature extraction function on an interval of values of one temporal variable (e.g., *the mean of the fever during an interval of days is greater than a value* v); finally, one is able to express a property of a (set) of multivariate time series as an interval temporal logic formula (e.g., *it is always true that during an interval in which the mean of the fever is greater than* v *there exists an interval in which the headache reaches a maximum value of pain of* v'.

In [8, 9], among others, a model for temporal decision trees with Halpern and Shoham's *Modal Logic for Time Intervals* (*HS*) [10] has been designed and used to learn patterns from datasets of multivariate time series. The logic HS is an unary propositional modal logic whose syntax encompasses all Boolean connectives plus one unary modal operator for each Allen's relation between two intervals, such as *during* or *later*.

In special cases, multivariate time series may display a richer internal structure than the one that can be captured with propositional HS. Two prominent examples of this situation are audio and electroencephalogram signals. In both cases, pre-processing of the original signal (in the first case, the single sound power expressed in dB, in the second case the electric power of each electrode expressed in mV) produces several temporal variables (in the first case audio frequencies, in the second one electric frequencies) which are naturally ordered and not mutually independent; we call such temporal variables *adjacent*. In this paper, we propose a simple first-order extension of HS that allows us to capture patterns of temporal datasets of multivariate time series with adjacent variables, by encompassing the possibility of comparing the natural order between temporal variables as well as the possibility of relating the behaviour of two variables in an interval of time.

2. First-Order Interval Temporal Logic for Adjacent Variables

The language of the *First-Order Modal Logic for Time Intervals for Adjacent Variables* (*FOHSa*, for short) encompasses a numerable set of first-order variables $X_1, X_2, \ldots$, a set of *unary function symbols* $f_1, f_2, \ldots$, arbitrary real constants, the set of *comparison operators* $<, \leq, =, \geq$, and $>$, standard Boolean operators, and unary *interval temporal logic operators*, one for each Allen's relation between any two intervals on a linear order, namely $\langle A \rangle$ (*meets*), $\langle L \rangle$ (*later*), $\langle O \rangle$

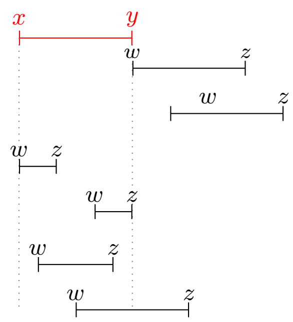
Modality	Definition	Example
$\langle A \rangle$ (after)	$[x,y]R_A[w,z] \quad \Leftrightarrow \quad y = w$	
$\langle L \rangle$ (later)	$[x,y]R_L[w,z] \quad \Leftrightarrow \quad y < w$	
$\langle B \rangle$ (begins)	$[x,y]R_B[w,z] \quad \Leftrightarrow \quad x = w \wedge z < y$	
$\langle E \rangle$ (ends)	$[x,y]R_E[w,z] \quad \Leftrightarrow \quad y = z \wedge x < w$	
$\langle D \rangle$ (during)	$[x,y]R_D[w,z] \quad \Leftrightarrow \quad x < w \wedge z < y$	
$\langle O \rangle$ (overlaps)	$[x,y]R_O[w,z] \quad \Leftrightarrow \quad x < w < y < z$	

Table 1
Allen's interval relations and (FO)HS modalities.

(*overlaps*), $\langle D \rangle$ (*during*), $\langle B \rangle$ (*begins*), $\langle E \rangle$ (*ends*), and their inverse ones (if $\langle Op \rangle$ is an interval operator, $\langle \overline{Op} \rangle$ denotes its inverse one). Formulas of FOHSa are built using the following grammar:

$$\begin{array}{lcl} t_f & ::= & f(V) \mid f(X) \mid v \\ At & ::= & t_f \bowtie t \mid V \bowtie V \\ \varphi & ::= & At \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle Op \rangle \varphi \mid \forall X \varphi, \end{array}$$

where V is a temporal variable, X is a first-order variable, f is a function, $\bowtie \in \{<, \leq, =, \geq, >\}$, $v \in \mathbb{R}$, and $Op \in \{A, L, B, E, D, O, \overline{A}, \overline{L}, \overline{B}, \overline{E}, \overline{D}, \overline{O}\}$.

Intuitively, a FOHSa formula is interpreted on a multivariate time series; towards a precise definition of the truth relation, though, a few definitions are necessary. First, a multivariate time series $T = \{V_1, \ldots, V_n\}$ is said to be *adjacent variables* if the set of time series $\{V_1, \ldots, V_n\}$ is linearly ordered by a relation $<_v$; the equality relation $=_v$ is defined in the standard way. Now, given the set $\mathbb{D}$ (the domain of all time series in T), we say that $\mathbb{I}(\mathbb{D})$ is the set of all *intervals* that can be built on $\mathbb{D}$, that is $\mathbb{I}(\mathbb{D}) = \{[x,y] \mid x, y \in \mathbb{D}, x <_t y\}$. Moreover, let $\mathcal{F} = \{F_1, \ldots, F_k\}$ be a set of *feature extraction function templates*, where each template F_i is, in turn, a set $F_i = \{f_i^j \mid j \in \mathbb{N}^+\}$, and each element of a set $f_i^j \in F_i$ is a function $f_i^j : \mathbb{R}^j \to \mathbb{R}$. A set F represents the interpretation of a function symbol f; since the value $f(V)$ is computed on an interval $[x,y]$, the specific function that must be used to compute it depends on the quantity $y - x + 1$; for example, if f is the (generalized) *mean* of a set of reals, computing the mean of V on the interval $[3,7]$ entails collecting the values $V(3), \ldots, V(7)$ and using the function *mean of 8 values*. Observe that we intentionally overload function symbols and their interpretation, as well as variable symbols and their interpretation, to ease the reading. A FOHSa *model* is a pair $\langle T, \mathcal{F} \rangle$, and the truth of a FOHSa formula φ on a model $\langle T, \mathcal{F} \rangle$ and an interval $[x,y]$ is given by the following clauses (see Tab. 1 for the semantics of the interval relations):

$$
\begin{array}{lll}
\langle T,\mathcal{F}\rangle,[x,y]\Vdash f_i(V_p)\bowtie f_j(V_q) & \text{iff} & f_i^{y-x+1}([V_p(x),V_p(x+1),\ldots,V_p(y)])\bowtie \\
 & & f_j^{y-x+1}([V_p(x),V_p(x+1),\ldots,V_p(y)]) \\
\langle T,\mathcal{F}\rangle,[x,y]\Vdash V_p\bowtie V_q & \text{iff} & V_p\bowtie V_q \\
\langle T,\mathcal{F}\rangle,[x,y]\Vdash \neg\varphi & \text{iff} & \langle T,\mathcal{F}\rangle,[x,y]\nVdash \varphi \\
\langle T,\mathcal{F}\rangle,[x,y]\Vdash \varphi\vee\psi & \text{iff} & \langle T,\mathcal{F}\rangle,[x,y]\Vdash \varphi \text{ or } \langle T,\mathcal{F}\rangle,[x,y]\Vdash \psi \\
\langle T,\mathcal{F}\rangle,[x,y]\Vdash \langle Op\rangle\varphi & \text{iff} & \text{there exists}[w,z] \text{ s.t. } [x,y]R_{Op}[w,x] \text{ and} \\
 & & \langle T,\mathcal{F}\rangle,[w,z]\Vdash \varphi \\
\langle T,\mathcal{F}\rangle,[x,y]\Vdash \forall X\varphi & \text{iff} & \text{for every variable } V \text{ it is the case that} \\
 & & \langle T,\mathcal{F}\rangle,[x,y]\Vdash \varphi[X/V],
\end{array}
$$

where $\bowtie\in\{<,\leq,=,\geq,>\}$.

A typical *audio signal* presents a spectrum of frequencies that range roughly from $20Hz$ to 20,000Hz. In audio signal processing of a sample, it is customary to extract its *spectral* representation, facilitating their interpretation in terms of audio frequencies. To this end, the most widespread adopted technique is known as extracting the *Mel-Frequency Cepstral Coefficients* (*MFCC*) [11]; obviously, MFCC includes a Fast Fourier Transform (FFT) step. As a result of several MFCC processing steps, a single sample is finally represented as a multivariate time series whose variables contain the value, at each time point, of the power of the signal at a specific frequency; frequencies are naturally ordered from the lowest one to the highest one. As another example, consider the signal recorded from an *electroencephalogram* executed on a (human) brain. In the most typical presentation, such a signal is the collection of the recording of several electrodes in a period of time. Again, each signal from an electrode present a spectrum of frequencies, usually from 0.5Hz to 50Hz. Again, a single sample is processed via FFT, applied to each signal of each electrode, resulting in the sample being represented as a multivariate time series whose variables contain the value of the electric power of the signal of a specific electrode at a specific frequency, and in this case as well frequencies are ordered from the lowest one to the highest one.

In time series processing, a set of *feature extraction functions* can be identified from the current literature (in most cases features have been presented as specific to a problem, but they often re-occur in different application areas). Among many others, examples of properties that can be expressed in FOHSa in the above domains include: *there exists a frequency for which it is always true that, during every interval when its mean value is lower than v there is an interval in which its maximum value is greater than v'* and *for every frequency it is true that if in an interval its mean value is lower than v then in some future interval there is a higher frequency whose mean value is greater than v'*.

3. Conclusions

We presented FOHSa, a novel logical system specifically designed to express non-propositional temporal interval properties that could be of interest in several applicative areas and different contexts including, for example, symbolic machine learning.

References

[1] P. Clark, T. Niblett, The CN2 Induction Algorithm, Machine Learning 3 (1989) 261–283.

[2] R. Rivest, Learning Decision Lists, Machine Learning 2 (1987) 229–246.

[3] L. Breiman, Bagging predictors, Machine Learning 24 (1996) 123–140.

[4] L. Breiman, Random forests, Machine Learning 45 (2001) 5–32.

[5] M. Kearns, L. Valiant, Cryptographic Limitations on Learning Boolean Formulae and Finite Automata, Journal of the ACM 41 (1994) 67–95.

[6] T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System, in: Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2016, pp. 785–794.

[7] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T. Liu, LightGBM: A Highly Efficient Gradient Boosting Decision Tree, in: Proc. of the 31st Advances in Neural Information Processing Systems (NIPS), 2017, pp. 3146–3154.

[8] G. Sciavicco, I. Stan, Knowledge extraction with interval temporal logic decision trees, in: Proc. of the 27th International Symposium on Temporal Representation and Reasoning (TIME), volume 178 of *LIPIcs*, 2020, pp. 9:1–9:16.

[9] G. Bechini, E. Losi, L. Manservigi, G. Pagliarini, G. Sciavicco, I. Stan, M. Venturini, Statistical rule extraction for gas turbine trip prediction, Journal of Engineering for Gas Turbines and Power (2023) 1 – 10.

[10] J. Halpern, Y. Shoham, A Propositional Modal Logic of Time Intervals, Journal of the ACM 38 (1991) 935–962.

[11] S. B. Davis, P. Mermelstein, Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences, IEEE Transactions on Acoustics, Speech and Signal Processing 28 (1980) 357–366.