

Multi-Frame Modal Symbolic Learning

Giovanni Pagliarini^{1,2}, Guido Sciavicco¹ and Ionel Eduard Stan^{1,2}

¹University of Ferrara, Italy

²University of Parma, Italy

Abstract

Symbolic learning is the sub-field of machine learning that deals with symbolic algorithms and models, which have been known for decades and successfully applied to a variety of contexts. The main limitation of symbolic models is the fact that they are essentially based on classical propositional logic, which implies that data with an implicit dimensional component, such as temporal (e.g., time series) or spatial data (e.g., images), cannot be properly dealt with within the standard symbolic framework. Recently, modal symbolic learning models have been proposed as a natural extension of classical ones to naturally deal with dimensional data, and successfully applied to temporal and spatial data. In this paper, we discuss the possibility of further extending such learning models to deal with multi-frame dimensional data, to be able to natively learn from instances represented by more than one dimensional description.

Keywords

Dimensional Data, Modal Logics, Modal Symbolic Learning


1. Introduction


The most iconic and fundamental separation between sub-fields of machine learning is the one between *functional* and *symbolic* learning. Functional learning is the process of learning a *function* that represents the theory underlying a certain phenomenon. Symbolic learning, on the other hand, is the process of learning a *logical description* that represents that phenomenon. Whether one or the other approach should be preferred raised a long-standing debate among experts, rooted in the fact that functional methods tend to be more versatile and statistically accurate than symbolic ones, while symbolic methods are able to extract models that can be interpreted, explained, and enhanced using human knowledge. From a logical standpoint, classical symbolic learning schemata are all characterized by the use of propositional logic (they are, in fact, sometimes called *propositional* methods), and can be classified along three main directions: the structure of the models (from strongly structured ones, such as decision trees [1, 2], to strongly unstructured ones, such as sets of independent rules [3]), the type of logic (crisp versus fuzzy [4, 5]), and the type of learning method (from purely deterministic to purely randomized). *Dimensional* data, such as temporal or spatial data, cannot be dealt with in a native way using propositional methods. Examples of naturally dimensional data include temporal

OVERLAY 2021: 3rd Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, September 22, 2021, Padova, Italy

✉ giovanni.pagliarini@unife.it (G. Pagliarini); guido.sciavicco@unife.it (G. Sciavicco); ioneleduard.stan@unife.it (I. E. Stan)

ORCID 0000-0002-8403-3250 (G. Pagliarini); 0000-0002-9221-879X (G. Sciavicco); 0000-0001-9260-102X (I. E. Stan)

 © 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

histories of patients, or objects described by spatial images; in all such cases, the dimensional component is usually implicit. The general to-go strategy to treat dimensional data with propositional models such as decision-trees is to *flatten* the dimensional component, effectively hiding it. This consists in massaging the data set in such a way that new variables are created for each dimensional variable A , which contain the values of A at different times, spatial locations, and so on; so for example, an instance that consists of a single-variate time series A with N (ordered) points ends up being represented as the (unordered) collection $A(1), A(2), \dots, A(N)$. Such a representation is called *lagged* in the temporal case, and *flattened* in the spatial one, and it allows one to use off-the-shelf propositional methods for the learning phase. Recently a new line in symbolic learning has emerged, in which classical symbolic algorithms are enhanced so that dimensional variables can be dealt by leveraging more expressive reasoning capabilities. To this end, propositional logic is replaced with propositional modal logic [6] in the learning schema, allowing one to natively express the relationships that emerge among the different worlds that describe each instance (e.g., time points, time intervals, extended areas, etc.). Modal logic can be declined into more practical languages, such as temporal or spatial logics, without loosing its basic principles, and the definition of modal symbolic learning schema immediately becomes a definition of a temporal, spatial, or spatial-temporal one. Modal decision trees, for example, have been studied in the temporal case in [7], and applied to real data in [8]. Dimensional data, however, is more complex than what can be captured by a single description. The temporal histories of patients, for instance, should be paired with patients' static data; the spatial descriptions of objects, as a different example, may require images from different angles; more in general, the instances of real-world data sets are often very complex, and require learning methods that can deal with such complexity.

In this paper we pave the way to a generalization of modal symbolic methods to the multi-frame case, and present some practical cases in which such a complex schema can be useful.

2. Multi-Frame Modal Symbolic Learning

A *multi-frame dimensional data set* \mathcal{I} is a finite collection of *instances* I_1, \dots, I_m , each of which is associated with (i.e., described by) r Kripke models M_1, \dots, M_r . In turn, in each model $M_j = (W_j, R_1^j, \dots, R_{s_j}^j, V_j)$, each world is characterized by the value of n_j distinct *attributes* $A_1^j, \dots, A_{n_j}^j$; as in the classical propositional case, the attributes define the propositional alphabet. We say that \mathcal{I} is *labeled* if it is partitioned in a finite number of classes $\mathcal{C} = \{C_1, \dots, C_k\}$. In other words, a single instance is described by more-than-one dimensional information. We assume that the number and the type of descriptions are consistent among instances; so, for example, a multi-frame dimensional data set may contain instances described by three frames each. To a multi-frame dimensional data set with r frames we associate r (in general, distinct) modal languages. The j -th language is a unary modal logic with s_j existential modalities $\diamond_1, \dots, \diamond_{s_j}$, and their corresponding universal versions $\square_1, \dots, \square_{s_j}$. Modalities are interpreted by the relations $R_1^j, \dots, R_{s_j}^j$, so that:

$$\begin{aligned} M_j, w \Vdash p & \quad \text{iff} \quad p \in V_j(w); \\ M_j, w \Vdash \neg\varphi & \quad \text{iff} \quad M_j, w \not\Vdash \varphi; \end{aligned}$$

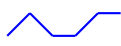
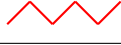




$$\begin{aligned}
M_j, w \Vdash \varphi \vee \psi & \text{ iff } M_j, w \Vdash \varphi \text{ or } M_j, w \Vdash \psi; \\
M_j, w \Vdash \diamond_t^j \varphi & \text{ iff } \exists v \text{ s.t. } wR_t^j v \text{ and } M_j, v \Vdash \varphi,
\end{aligned}$$

where \diamond_t^j is the t -th modality of the j -th language. Multi-frame modal symbolic learning consists of enhancing modal symbolic learning methods with the possibility of learning from more-than-one dimensional description at the same time, and describing the learned knowledge using the correct logic. Multi-frame dimensional data sets capture dimensional situations quite naturally; but just as it happens in modal symbolic learning, in the multi-frame case too we need to *concretize* the learning models and the associated languages to specific modal logics to adapt them to the real-world cases. Dimensional data is generally represented in an implicit form. This means that, for example, time series and images are usually linearized, and expressed as sets of numbers; different, but equivalent, representations can be used to capture dimensional situations (*by rows* and *by columns* are just two very popular examples). A key observation is that the information in real data, in general, is not point-based (think, for example, to time series or images: the values of single time points or single pixels are not really informative). One around this problem is to employ modal logics in which worlds do not correspond to single points, but to sets of points. Halpern and Shoham's interval temporal logic (HS) [9] allows one to express properties of intervals in the temporal case. By generalizing HS to any number of dimensions, we obtain a family of logics that we can denote by HS^d , where $d \in \mathbb{N}$; HS^0 is just propositional logic, HS^1 is the original Halpern and Shoham logic of intervals, and HS^2 is the natural logical generalization of Rectangle Algebra [10]. In HS^d , worlds are *hyperrectangles* with edges parallel to the axis, and connected by the d -dimensional generalizations of Allen's interval relations. In this way, a set of propositional letters emerges naturally. Fixed a frame, a world, and an attribute A , these are of the type:

$$f(A) \bowtie_{\sim\gamma} a$$

where a is a value of the domain of A , f is a function, $\gamma \in (0, 1] \subset \mathbb{R}$, and $\bowtie \in \{<, \leq, =, \geq, >\}$, and $\sim \in \{<, \leq, \geq, >\}$. Propositional letters of this type are interpreted over hyperrectangles; for example, if $d = 1$ (that is, data is one-dimensional, e.g., temporal), f is the identity function, \bowtie is $<$, \sim is $>$, γ is 0.8, A is the temperature, and $a = 37.5$, then $A <_{>0.8} a$ represents the proposition *more than the 80% of the values of the temperature during the current interval are below 37.5*. By varying the function f , one can produce more complex assertions on hyperrectangles.

For a better understanding of how multi-frame modal symbolic learning can be useful in practical case, consider, again, the medical example, as in Fig. 1 (top). In this situation, we have a three-frame data set. The first frame is static, and it is associated to propositional logic (HS^0); the second frame is one-dimensional, and, specifically, temporal: we associate it with HS^1 , so that the information concerning the temporal attributes is learned by intervals; the third frame is two-dimensional, that is, spatial: by associating it to HS^2 , we can learn patterns of rectangles and colors in them. Instances are labeled, making this data set suitable for knowledge extraction by classification. One can observe how data are represented in concrete form (Fig. 1, bottom left), and how we can extract knowledge in symbolic form from them (Fig. 1, bottom, right), with modal decision trees or sets of modal rules in HS^d .

Patient/Instance	Static data	Temporal data	Image (R,G,B)	Class
P_1	Age (A): 37, Weight (W): 70 Gender (G): M	Temp (T):  Press (P): 		C_1
P_2	Age (A): 49, Weight (W): 81 Gender (G): M	Temp (T):  Press (P): 		C_2
...

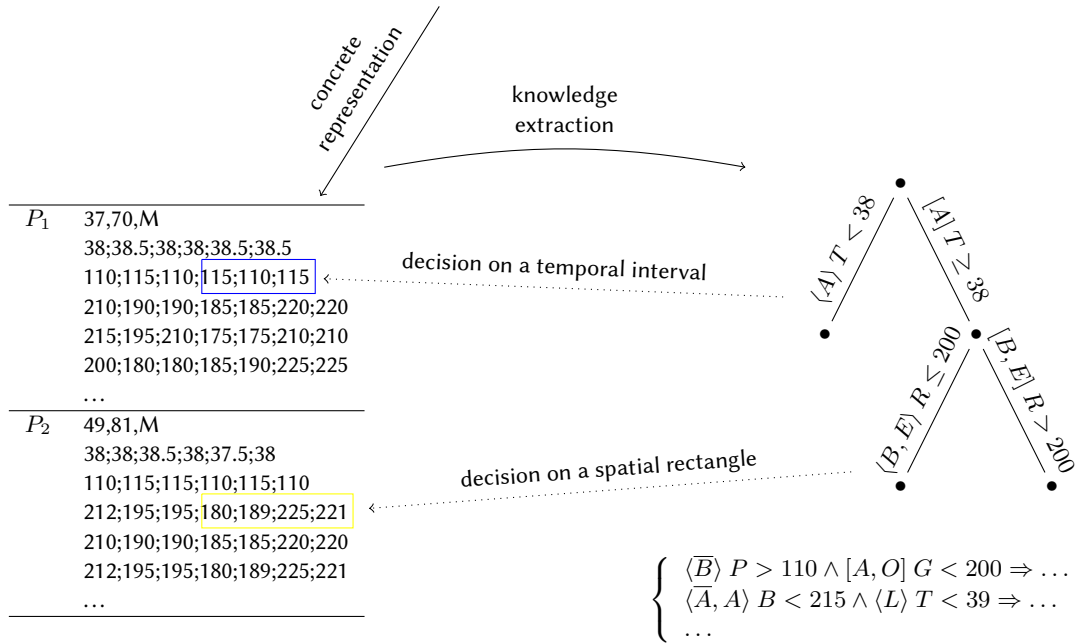


Figure 1: Example of multi-frame dimensional data set.

3. Conclusions

In this paper we paved the way towards multi-frame modal symbolic learning. We defined multi-frame dimensional data sets, the learning framework, and described how to concretize it to deal with implicit dimensional data in an intuitive way.

Acknowledgments

We thank the INdAM GNCS 2020 project *Strategic Reasoning and Automated Synthesis of Multi-Agent Systems* for partial support.

References

- [1] J. R. Quinlan, Induction of decision trees, *Machine Learning* 1 (1986) 81–106.
- [2] J. R. Quinlan, Simplifying decision trees, *International Journal of Human-Computer Studies* 51 (1999) 497–510.
- [3] A. K. H. Tung, *Rule-based Classification*, Springer, 2009, pp. 2459–2462.
- [4] S. Kundu, Similarity relations, fuzzy linear orders, and fuzzy partial orders, *Fuzzy Sets and Systems* 109 (2000) 419–428.
- [5] S. Ovchinnikov, Similarity relations, fuzzy partitions, and fuzzy orderings, *Fuzzy Sets and Systems* 40 (1991) 107 – 126.
- [6] P. Blackburn, M. d. Rijke, Y. Venema, *Modal Logic*, Cambridge University Press, 2001.
- [7] G. Sciavicco, I. E. Stan, Knowledge Extraction with Interval Temporal Logic Decision Trees, in: *Proc. of the 27th International Symposium on Temporal Representation and Reasoning*, volume 178 of *LIPICs*, 2020, pp. 9:1–9:16.
- [8] F. Manzella, G. Pagliarini, G. Sciavicco, I. E. Stan, Interval Temporal Random Forests with an Application to COVID-19 Diagnosis, in: *Proc. of the 28th International Symposium on Temporal Representation and Reasoning*, volume 206 of *LIPICs*, 2021, pp. 3:1–3:17.
- [9] J. Y. Halpern, Y. Shoham, A propositional modal logic of time intervals, *Journal of the ACM* 38 (1991) 935–962.
- [10] P. Balbiani, J. Condotta, L. Fariñas del Cerro, A model for reasoning about bidimensional temporal relations, in: *Proc. of the 6th International Conference on Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann, 1998, pp. 124–130.