Learning Temporal Logic Formulas from Time-Series Data

Laura Nenzi
Department of Engineering and Architecture, University of Trieste, Italy

Abstract

In this talk, we provide an overview of recent advancements in the field of mining formal specifications from time-series data, with a specific focus on learning Signal Temporal Logic (STL) formulae.

2012 ACM Subject Classification Theory of computation → Modal and temporal logics

Keywords and phrases Temporal Logic, Mining Specifications

Digital Object Identifier 10.4230/LIPIcs.TIME.2023.1

Category Invited Talk

Funding This research was partially supported by the Austrian FWF projects ZK-35, and by the consortium iNEST funded by PNRR – Missione 4 Componente 2, Investimento 1.5 – D.D. 1058 23/06/2022, ECS_00000043).

Acknowledgements The research surveyed in this talk is joint work with Luca Bortolussi, Eric Medvet, Jyotirmoy V. Deshmukh, Simone Silvetti, Federico Pigozzi, Patrick Indri, Sara Mohammadinejad, Ezio Bartocci, and Alberto Bartoli.

1 Extended Abstract

The abundance of available data has led to a significant increase in the utilization of machine learning techniques for describing and analyzing systems. Although these methods are adept at generating powerful black-box models and performing well in complex, high-dimensional scenarios, they often lack measures of uncertainty for individual estimates and fail to fully understand the underlying mechanisms driving the obtained outcomes. This limitation becomes particularly crucial when dealing with safety-critical systems like smart healthcare and self-driving cars, where failures can have severe consequences and incur substantial costs. Therefore, it is imperative for designers to gain a comprehensive understanding of the phenomena these systems capture and to extract interpretable information from the data they produce.

To tackle this challenge, recent research has explored the application of learning Temporal Logic (TL) formulae as a powerful approach for extracting human-interpretable information from data. TL is a formal language that offers precise specifications which are easily comprehensible to humans, allowing for the expression of complex system properties in a clear manner. Furthermore, it provides verification algorithms that can automatically evaluate the satisfaction of these properties. In this talk, we aim to provide a comprehensive overview of the latest advancements in this field. We focus in particular on the learning of Signal Temporal Logic (STL) formulae [2]. STL is a linear-time TL very suitable to describe properties associated with real-time trajectories.

First, we present a framework for a supervised learning scenario [4]. We focus on a two-label classification problem, specifically targeting the discrimination between regular and anomalous trajectories. The objective is to develop a technique that learns a Signal Temporal Logic (STL) formula capable of effectively distinguishing between the two sets of labeled data. The desired formula should be satisfied by the regular trajectories to the greatest extent possible, while not being satisfied by the anomalous ones. To accomplish this, the proposed methodology leverages genetic programming, a form of Evolutionary Computation (EC), to extract the structure of the formula, and Bayesian optimization to learn the formula's parameters.

Second, we explore a semi-supervised scenario, which poses a more complex problem as it considers datasets comprising solely regular behaviors [1]. We introduce an algorithm that utilizes the Grammar-Guided Genetic Programming (G3P) technique to learn an ensemble of STL formulas. This ensemble is then utilized for effectively detecting anomalous behaviors within the system.

Third, we explore the integration of spatial considerations within the study of system behavior. To achieve this, we introduce the Spatio-Temporal Reach and Escape Logic (STREL) [3], an extension of STL that incorporates a variety of spatial operators. This extension enables the modeling and analysis of spatial-temporal properties, providing a comprehensive framework for capturing complex behaviors that involve both temporal and spatial aspects. We present a method that leverages a Parametric STREL (PSTREL) for automatic feature extraction from the given spatio-temporal data. This is accomplished by projecting the data onto the parameter space of PSTREL. Using an agglomerative hierarchical clustering technique, we ensure the satisfaction of a distinct STREL formula in each cluster.

We demonstrate the versatility of these techniques through multiple case studies spanning various domains, including naval surveillance, train speed regulation, secure water treatment, urban transportation, and epidemiological analysis. We conclude the talk by discussing the remaining challenges and future prospects, while also providing an overview of the latest ongoing research.

References

1. Patrick Indri, Alberto Bartoli, Eric Medvet, and Laura Nenzi. One-shot learning of ensembles of temporal logic formulas for anomaly detection in cyber-physical systems. In *Genetic Programming - 25th European Conference, EuroGP 2022, Held as Part of EvoStar 2022, Madrid, Spain, April 20-22, 2022, Proceedings*, pages 34–50. Springer, 2022. `doi:10.1007/978-3-031-02056-8_3`.
2. Oded Maler and Dejan Nickovic. Monitoring temporal properties of continuous signals. In Yassine Lakhnech and Sergio Yovine, editors, *Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems, Joint International Conferences on Formal Modelling and Analysis of Timed Systems, FORMATS 2004 and Formal Techniques in Real-Time and Fault-Tolerant Systems, FTRTFT 2004, Grenoble, France, September 22-24, 2004, Proceedings*, Lecture Notes in Computer Science, pages 152–166. Springer, 2004. `doi:10.1007/978-3-540-30206-3_12`.
3. Laura Nenzi, Ezio Bartocci, Luca Bortolussi, and Michele Loreti. A logic for monitoring dynamic networks of spatially-distributed cyber-physical systems. *Log. Methods Comput. Sci.*, 18(1), 2022. `doi:10.46298/lmcs-18(1:4)2022`.
4. Laura Nenzi, Simone Silvetti, Ezio Bartocci, and Luca Bortolussi. A robust genetic algorithm for learning temporal specifications from data. In *Quantitative Evaluation of Systems - 15th International Conference, QEST 2018, Beijing, China, September 4-7, 2018, Proceedings*, pages 323–338. Springer, 2018. `doi:10.1007/978-3-319-99154-2_20`.