**Workshop on**

# Algorithms & Theories for the Analysis of Event Data (ATAED'2022)

**Workshop, June 24, 2022**

supported by the IEEE Task Force on Process Mining

Satellite event of the conference

**43rd International Conference on Application and Theory of Petri Nets and Concurrency (Petri Nets 2022)**

.

# Preface

Ehrenfeucht and Rozenberg defined regions more than 30 years ago as sets of nodes of a finite transition system. Every region relates to potential conditions that enable or disable transition occurrences in an associated elementary net system. Later, similar concepts were used to define regions for Petri nets from languages as well. Both *state-based* and *language-based* approaches aim to constrain a Petri net by adding places deduced from the set of *region*s. By now, many variations have been proposed, e.g., approaches dealing with multiple tokens in a place, region definitions for Petri nets with inhibitor arcs, extensions to partial languages, regions for infinite languages, etc.

Initially, region theory focused on *synthesis*. We require the input and the behavior of the resulting Petri net to be equivalent. Recently, region-based research started to focus on *process mining* as well where the goal is *not* to create an equivalent model but to *infer* new knowledge from the input. Process mining examines observed behavior rather than assuming a complete description in terms of a transition system or prefix-closed language. For this reason, one needs to deal with new problems such as noise and incompleteness. Equivalence notions are replaced by trade-offs between fitness, simplicity, precision, and generalization. A model with good *fitness* allows for most of the behavior seen in the event log. A model that does not *generalize* is "overfitting". Overfitting is the problem that a very specific model is generated whereas it is obvious that the log only holds example behavior. A model that allows for "too much behavior" lacks *precision*. Simplicity is related to Occam's Razor which states that "one should not increase, beyond what is necessary, the number of entities required to explain anything". Following this principle, we look for the *simplest* process model that can explain what was observed in the event log. Process discovery from event logs is very challenging because of these and many other trade-offs. Clearly, there are many theoretical process-mining challenges with a high practical relevance that need to be addressed urgently.

All these challenges and opportunities are the motivation for organizing the *Algorithms & Theories for the Analysis of Event Data* (ATAED) workshop. The workshop first took place in Brussels in 2015 as a succession of the *Applications of Region Theory* (ART) workshop series. From there on, the workshop moved to Toruń (2016), Zaragoza (2017), Bratislava (2018), Aachen (2019), and virtually in 2020 (due to the COVID-19 pandemic). After the success of these workshops, it is only natural to bring together researchers working on region-based synthesis and process mining again.

The ATAED'2022 workshop took place as a physical workshop on June 21st, 2022 and was a satellite event of the 43rd International Conference on Application and Theory of Petri Nets and Concurrency (Petri Nets 2022), held in Bergen, Norway.

Papers related to process mining, region theory and other synthesis techniques were presented at the ATAED'2022, divided over three content-oriented sessions, i.e., "Stochastics & Statistics", "Region Theory" and "Strategies for Behavioral Analysis". All the techniques presented have in common that "lower-

level" behavioral descriptions (event logs, partial languages, transition systems, etc.) are used to create "higher level" process models (e.g., various classes of Petri nets, BPMN, or UML activity diagrams). In fact, all techniques that aim at learning or checking concurrent behavior from transition systems, runs, or event logs were welcomed. The workshop was supported by the IEEE Task Force on Process Mining (`www.tf-pm.org/`).

After a careful reviewing process, six papers (out of a total of ten submissions) were accepted for the workshop. We thank the reviewers for providing the authors with valuable and constructive feedback. We thank the authors and the presenters for their wonderful contributions.

Enjoy reading the proceedings!

Robert Lorenz, Jan Martijn van der Werf, and Sebastiaan J. van Zelst
June 2022

### Program committee of ATAED'2022

Abel Armas Cervantes, QUT, Australia
Luca Bernardinello, Universitá degli studi di Milano-Bicocca, Italy
Paolo Ceravolo, University of Milan, Italy
Jochen De Weerdt, KU Leuven, Belgium
Benoît Depaire, Hasselt University, Belgium
Jörg Desel, FernUni Hagen, Germany
Claudio Di Ciccio, Sapienza University of Rome, Italy
Chiara Di Francescomarino, FBK-IRST, Italy
Dirk Fahland, TU Eindhoven, The Netherlands
Stefan Haar, LSV CNRS & ENS de Cachan, France
Anna Kalenkova, University Adelaide, Australia
Jetty Kleijn, Leiden University, The Netherlands
Sander Leemans, RWTH Aachen University, Germany
Robert Lorenz, University of Augsburg, Germany (co-chair)
Lisa Mannel, RWTH Aachen University, Germany
Marta Pietkiewicz-Koutny, Newcastle University, United Kingdom
Andrey Rivkin, Free University of Bozen Bolzano, Italy
Daniel Schuster, Fraunhofer FIT/RWTH Aachen University, Germany
Arik Senderovich, York University, Canada
Ronny Tredup, University of Rostock, Germany
Lijie Wen, Tsinghua University, China
Alex Yakovlev, Newcastle University, United Kingdom
Jan Martijn van der Werf, Utrecht University, The Netherlands (co-chair)
Sebastiaan J. van Zelst, Fraunhofer FIT/RWTH Aachen University, Germany (co-chair)