Workshop Proceedings

Workshop on

Algorithms & Theories for the Analysis of Event Data (ATAED'2017)

Zaragoza, Spain, June 26-27, 2017

Satellite event of the conferences

17th International Conference on Application of Concurrency to System Design (ACSD 2017)

38th International Conference on Application and Theory of Petri Nets and Concurrency (PN 2017)

Edited by
Wil van der Aalst, Robin Bergenthum, and Josep Carmona

.

Preface

Ehrenfeucht and Rozenberg defined regions about 25 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 *regions*. 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. The second workshop took place in Toruń in 2016. After the success of both workshops, it is only natural to bring together researchers working on region-based synthesis and process mining again.

The ATAED'2017 workshop took place in Zaragoza on June 26-27, 2017 and was a satellite event of both the 38th International Conference on Application and Theory of Petri Nets and Concurrency (Petri Nets 2017) and the 17th International Conference on Application of Concurrency to System Design (ACSD 2017). Papers related to process mining, region theory and other synthesis techniques were presented at ATAED'2017. These techniques 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.win.tue.nl/ieeetfpm/).

After a careful reviewing process, twelve papers were accepted for the workshop. Overall, the quality of the submitted papers was good and most submissions matched the workshop goals very well. We thank the reviewers for providing the authors with valuable and constructive feedback. Moreover, we were honored that *Thomas Hildebrandt* was willing to give an invited talk on *"Modelling & Mining Event-based Concurrent Declarative Processes as Dynamic Condition Response Graphs"*. We thank Thomas, the authors, and the presenters for their wonderful contributions.

Enjoy reading the proceedings!

Wil van der Aalst, Robin Bergenthum, and Josep Carmona
June 2017

Program committee of ATAED'2017

Wil van der Aalst, TU Eindhoven, The Netherlands (co-chair)
Eric Badouel, INRIA Rennes, France
Robin Bergenthum, FernUni Hagen, Germany (co-chair)
Luca Bernardinello, Universitá degli studi di Milano-Bicocca, Italy
Seppe vanden Broucke, KU Leuven, Belgium
Andrea Burattin, University of Innsbruck, Austria
Josep Carmona, UPC Barcelona, Spain (co-chair)
Paolo Ceravolo, University of Milan, Italy
Claudio Di Ciccio, Vienna University of Economics and Business, Austria
Benoît Depaire, Hasselt University, Belgium
Jörg Desel, FernUni Hagen, Germany
Dirk Fahland, TU Eindhoven, The Netherlands
Diogo Ferreira, Universidade de Lisboa, Portugal
Luciano García-Bañuelos, University of Tartu, Estonia
Stefan Haar, LSV CNRS & ENS de Cachan, France
Gabriel Juhás, Slovak University of Technology, Slovak Republic
Anna Kalenkova, Higher School of Economics NRU, Russia
Jetty Kleijn, Leiden University, The Netherlands
Wen Lijie, Tsinghua University, China
Robert Lorenz, Uni Augsburg, Germany
Manuel Mucientes, University of Santiago de Compostela, Spain
Marta Pietkiewicz-Koutny, Newcastle University, GB
Hernán Ponce de León, fortiss GmbH, Germany
Matthias Weidlich, Humboldt-Universität zu Berlin, Germany
Jochen De Weerdt, KU Leuven, Belgium
Moe Wynn, Queensland University of Technology, Australia
Alex Yakovlev, Newcastle University, GB

Table of Contents