

Conformance Checking in Process Mining

Mieke Jans, Jochen De Weerd, Benoît Depaire, Marlon Dumas, Gert Janssenswillen

1. Introduction

In February 2018, a two-day brainstorming was organized by the Flemish Scientific Research Community on Process Mining¹ to discuss current challenges in the field of conformance checking. An international group of process mining researchers with interest and expertise in conformance checking, affiliated with Hasselt University, KU Leuven, the Polytechnic University of Catalonia, the Pontifical Catholic University of Chile, RWTH Aachen University, the Technical University of Eindhoven and Wirtschaftsuniversität Wien gathered for two days. The goal of this brainstorming was to have discussions on the future of conformance checking, articulating shortcomings of current techniques as well as opportunities and challenges for the field.

The conclusion of this brainstorming was that, on one hand, there is a widespread consensus about the problems and challenges the field is still facing, which will be described briefly in the next section. On the other hand, it was also established that there are still many different views on how these challenges should be tackled, as well as diverging views on the precise definition and characterization of what *conformance checking* is. As the diverging views are indicative of the relatively young nature of the field, and represent different research streams to approach the problem, an attempt to harmonize these diverging views into a single definition of conformance checking, can be considered unwise at this moment. It was therefore decided to initiate this Special Issue on conformance checking, in order to facilitate the further development of the field in an unconstrained manner.

¹<https://srcprocessmining.com/>

2. Conformance Checking in Process Mining and Associated Challenges

Process mining combines process model-driven approaches and data mining techniques to create methods and tools providing fact-based insights into processes and supporting process improvements (van der Aalst 2016). Over the past two decades, the field has matured significantly from an initial focus on discovering control-flow models from event data to a myriad of new research challenges (Recker and Mendling 2016). The growth of the field of process mining is reflected by the growing community of researchers which form a recognizable and constant presence at conferences such as CAiSE and BPM - and more recently by the birth of its own International Conference on Process Mining in 2019. It is also telling that industry has witnessed an uptake of process mining techniques resulting in a growing market approaching \$180 million (Gartner 2019).

One of the research challenges within process mining that has gained significant traction over the past decade is the development of conformance checking techniques. Conformance checking relates modelled with recorded process behaviour. It comprises techniques and methods to compare and analyse observed process behaviour in the presence of a process model. This Special Issue invited researchers in the field of conformance checking to submit original research papers that explore the current boundaries of the research domain. Submissions could introduce new paradigms, address promising application domains or tackle interesting challenges which have the potential to uncover many new research opportunities for the future. We envisioned some application domains and challenges of conformance checking research, which we list below.

Process diagnostics - Process diagnostics is the broad area of exploratory data analysis which provides value to the user by adding structure and context to the original process data. Deviation detection, root cause analysis, deviation categorisation and performance analysis are examples of specific application domains that provide promising research opportunities for conformance checking techniques.

Compliance - Compliance refers to the procedures and internal control systems that organisations have in place to comply with regulations and laws,

such as e.g. GDPR. Within the area of compliance, conformance checking can play an important role in checking process rules.

Process model improvement - The comparison of recorded and modelled process behaviour generates various opportunities towards process redesign and improvement. Conformance checking creates opportunities for process model improvement challenges, such as model enrichment which tries to visually enrich an existing process model, as well as model repair which alters process models to better reflect reality.

Process Discovery Quality Assessment - Process Discovery Quality Assessment is the area concerned with the evaluation of process discovery algorithms. One of the main challenges in this application domain relates to the availability of suitable quality metrics to assess the quality of models with respect to the data.

Predictive business process monitoring - Predictive Business Process Monitoring refers to the prediction of the future of incomplete process instances. Conformance checking can play an important role in this application field by providing both new input features as well as the actual target variable for such prediction problems.

Potential challenges with respect to conformance checking include, but are not limited to:

Representing uncertainty and preventing bias - When conformance checking is used to make claims to what extent the process model conforms the underlying process, or the other way around, rather than the observed data, it becomes important to realize that the observed data is only an incomplete sample of the possible behaviour. Consequently, the conformance measures have the potential to be imprecise or even biased. Research is needed to address these issues.

Computational feasibility and online conformance checking - Computational feasibility is an important challenge in the field of conformance checking. As the size of event logs persistently become larger, the conformance metrics need to become increasingly computationally efficient. As event logs can even become so large that it is no longer possible to store

it all, conformance checking needs to operate in an online setting, which presents its own set of challenges.

Desired properties of conformance measures - A substantial part of conformance checking deals with the development of measures to express process conformance in a quantitative way. Despite the various measures that have been presented over the past decade, many challenges remain open with respect to desired properties of such measures and to what extent current measures meet these requirements. Possible requirements, among others, are logical consistency, robustness, confidence and comparability of the measures.

Multi-perspective and multi-paradigm conformance checking - Where in the early days of process mining the control-flow perspective and a procedural paradigm expressed in Petri net notation dominated the field of process mining, different perspectives, paradigms and modelling notations have been explored over time. Currently, various techniques even exist which mix these perspectives, paradigms and notations, introducing a whole new set of challenges to conformance checking.

3. Contents of the Special Issue

In response to the call for papers for this special issue, overall nine submissions were received. After two rounds of rigorous peer-review and editorial review, we selected five papers for inclusion in the special issue. Each paper was reviewed by two expert reviewers plus one of the guest editors. The selected articles address both application domains and fundamental challenges of conformance checking.

The first two papers relate to the application of conformance checking for process diagnostics. The paper “Stochastic Process Mining: Earth Movers’ Stochastic Conformance” by Leemans, van der Aalst, Brockhoff, and Polyvyanyy proposes the use of stochastic conformance checking and provides a new approach to compare event logs and process models (Leemans et al. 2021). The paper argues that such a stochastic approach tackles traditional conformance measures and shows that it enables detailed diagnostics projected on both model and log. The paper of Boltenhagen, Chatain, and Carmona, titled “Model-Based Trace Variant Analysis of Event Logs” introduces an analysis framework for trace variants capable of taking into account a process

model (Boltenhagen, Chatain and Carmona 2020). By combining alignments with SAT encoding and using an intelligent sampling algorithm for large data sets, better subsets of traces can be formed, guided by a normative underlying process model. Additional to process diagnostics, this work also touches upon process model improvement, another application where conformance checking adds value.

The other three papers address fundamental research challenges of conformance checking algorithms. The paper “Conformance checking of mixed-paradigm process models”, by van Dongen, De Smedt, Di Ciccio, and Mendling, addresses the challenge of checking conformance in a multi-paradigm setting (Dongen et al. 2020). It presents the first mixed-paradigm process model conformance checking approach, for models that combine declarative and procedural approaches. The approach, illustrated on real-life logs, shows how event data can be replayed on procedural models, while at the same time keeping track of declarative constraints, in order to compute trace fitness.

The fourth and fifth paper –“Empowering conformance checking using Big Data through horizontal decomposition”, by Valencia-Parra, Varela-Vaca, Gómez-López, Carmona, and Bergenthum, and ”Orientation and conformance: A HMM-based approach to online conformance checking” by Lee, Burattin, Muñoz-Gama, and Sepulveda– are addressing the challenges of computational feasibility and online conformance checking (Lee et al. 2020; Valencia-Parra et al. 2021). The former proposes an innovative horizontal process model decomposition. The approach is illustrated with the well-known A* Alignments algorithms, as well as using the relatively new Constraint Programming Paradigm. The latter proposes a framework for online conformance checking based on Hidden Markov Models. The approach alternates between locating a running case in the process and computing conformance. The paper reports on experimental results, demonstrating the approach’s feasibility.

To sum up, we believe this collection of papers not only makes an interesting read on conformance checking but is also a nice showcase of how research can advance both the algorithms and possible applications of conformance checking. We hope the listed application domains and research challenges, although not exhaustive, might serve as an anchor point for positioning future research. Additionally, we encourage other researchers to

continue investigating this research path. But for now, sit down, relax, and enjoy your read.

References

- Boltenhagen, Mathilde, Thomas Chatain and Josep Carmona (2020). “Model-based trace variant analysis of event logs”. In: *Information Systems*, p. 101675. ISSN: 0306-4379.
- Dongen, Boudewijn F. van et al. (2020). “Conformance checking of mixed-paradigm process models”. In: *Information Systems*, p. 101685.
- Gartner (June 2019). *Market Guide for Process Mining*. Tech. rep.
- Lee, Wai Lam Jonathan et al. (2020). “Orientation and conformance: A HMM-based approach to online conformance checking”. In: *Information Systems*, p. 101674. ISSN: 0306-4379. DOI: <https://doi.org/10.1016/j.is.2020.101674>. URL: <https://www.sciencedirect.com/science/article/pii/S0306437920301290>.
- Leemans, Sander J. J. et al. (Feb. 2021). “Stochastic process mining: Earth movers’ stochastic conformance”. en. In: *Information Systems*, p. 101724. ISSN: 0306-4379. DOI: [10.1016/j.is.2021.101724](https://doi.org/10.1016/j.is.2021.101724). URL: <https://www.sciencedirect.com/science/article/pii/S0306437921000041> (visited on 30/03/2021).
- Recker, Jan and Jan Mendling (2016). “The state of the art of business process management research as published in the BPM Conference”. In: *Business & Information Systems Engineering* 58.1, pp. 55–72. (Visited on 31/03/2016).
- Valencia-Parra, Álvaro et al. (2021). “Empowering conformance checking using Big Data through horizontal decomposition”. In: *Information Systems* 99, p. 101731. ISSN: 0306-4379. DOI: <https://doi.org/10.1016/j.is.2021.101731>. URL: <https://www.sciencedirect.com/science/article/pii/S0306437921000077>.
- van der Aalst, Wil (2016). *Process Mining: Data Science in Action*. Springer.