Monotone Precision and Recall Measures for Comparing Executions and Specifications of Dynamic Systems

(Extended Abstract)[1]

Artem Polyvyanyy[2], Andreas Solti[3], Matthias Weidlich[4],
Claudio Di Ciccio[5], Jan Mendling[6]

1 Motivation

The behavioural comparison of systems is an important concern of software engineering research. For example, the areas of specification discovery and specification mining are concerned with measuring the consistency between a collection of execution traces and a program specification. This problem is also tackled in process mining with the help of measures that describe the quality of a process specification automatically discovered from execution logs.

Technically, behavioural comparisons are often formulated in a *relative* manner, defining a *quotient* of some aspect of one behaviour over the same aspect of another behaviour. For instance, the quotients of the behaviours of a system at different points in time reveal how the system has changed. In process mining, in turn, the quotient of the behaviour of a system as recorded in a log over the behaviour as specified can be used to analyse the trustworthiness of the latter [Ca18]. Yet, defining such quotients is challenging: A recent commentary on measures in process mining identifies a set of intuitive properties and shows that none of the available measures fulfils them [Ta18]. That is, measures lack monotonicity or cannot handle infinite behaviour.

[1] The original article appeared in the ACM Transactions on Software Engineering and Methodology (TOSEM) [Po20]

[2] The University of Melbourne, Australia, artem.polyvyanyy@unimelb.edu.au
[3] Vienna University of Economics and Business, Austria, solti@ai.wu.ac.at
[4] Humbold-Universität zu Berlin, Germany, matthias.weidlich@hu-berlin.de
[5] Sapienza University of Rome, Italy, diciccio@di.uniroma1.it
[6] Vienna University of Economics and Business, Austria, mendling@ai.wu.ac.at

2 Approach

In this work, we address the problem of behavioural comparison by introducing a new framework for the definition of behavioural quotients. To this end, we approach the problem of behavioural comparison based on the notion of a formal language, which is a suitable starting point to capture the sequential (state-based) behaviour of a dynamic system. Specifically, we address the problem of *how to define meaningful quotients for behavioural comparison of finite and infinite languages.*

Our contributions include a framework for the definition of behavioural quotients that guarantee desired properties. Our framework is then used to instantiate two quotients that are grounded in the cardinality of a language (for finite languages) and the entropy of an automaton (for finite and infinite languages). It was recently shown that our entropy-based measures satisfy all the properties put-forward in the literature for recall and precision measures [STvdA19]. We further show how the proposed quotients can be used to define monotone precision and recall measures between the behaviour as recorded in an execution log of a system and the behaviour captured in a specification of the system.

3 Evaluation

Based on the jBPT library [PW13], we implemented the proposed precision and recall quotients in a tool that is publicly available.[7] Comprehensive experiments with execution logs of real IT systems highlight that existing measures indeed violate the monotonicity property. Moreover, the results of controlled experiments underpin that our quotients enable meaningful conclusions on the relation of the behaviours of two systems. Finally, we explore the scalability of the computation of our measures. The results illustrate the importance of convergence properties on the runtime of the method.

Bibliography

[Ca18] Carmona, Josep; Van Dongen, Boudewijn F.; Solti, Adreas; Weidlich, Matthias: Conformance Checking – Relating Processes and Models. Springer, 2018.

[Po20] Polyvyanyy, Artem; Solti, Andreas; Weidlich, Matthias; Ciccio, Claudio Di; Mendling, Jan: Monotone Precision and Recall Measures for Comparing Executions and Specifications of Dynamic Systems. ACM Trans. Softw. Eng. Methodol., 2020. In press.

[PW13] Polyvyanyy, Artem; Weidlich, Matthias: Towards a Compendium of Process Technologies - The jBPT Library for Process Model Analysis. CAiSE'13 Forum. Vol. 998 of CEUR Workshop Proceedings. CEUR-WS.org, pp. 106–113, 2013.

[STvdA19] Syring, Anja F.; Tax, Niek; van der Aalst, Wil M. P.: Evaluating Conformance Measures in Process Mining Using Conformance Propositions. TOPNOC 14:192–221, 2019.

[Ta18] Tax, Niek; Lu, Xixi; Sidorova, Natalia; Fahland, Dirk; van der Aalst, Wil M.P.: The imprecisions of precision measures in process mining. IPL, 135:1–8, 2018.

[7] `https://github.com/jbpt/codebase/tree/master/jbpt-pm`