# PHD Course on:

# Achieving Business Process Intelligence through Process Mining

***State of the Art and Recent Research Advances***

**Credits:** 3 CFU (20 academic hours)

**Lecturers:**

- Fabrizio Maria Maggi (University of Tartu, Estonia)
- Andrea Marrella (Sapienza, University of Rome)

## Short Description

The course will provide an introduction to the key analysis techniques in process mining that allow users to automatically learn process models from raw event logs (process discovery) and to check if reality, as recorded in the event logs, conforms to the models and vice versa (conformance checking).

In the context of process discovery and conformance checking, well-established techniques available in the literature will be presented, as well as more recent advances that are currently investigated in the process mining community, such as process mining with declarative specifications or BPMN. Other branches of the process mining field will also be presented like performance analysis, bottleneck detection, social network analysis.

In addition, the course will discuss how recent techniques from academic artificial intelligence research can be employed to advance the state-of-the art in process mining. Finally the course will provide easy-to-use software, real-life data sets, and practical skills to directly apply the theory in a variety of application domains.

## Overview

Process mining provides a new perspective to improve business processes in a variety of application domains where there is the need to improve the workflows performance (e.g., reducing costs and production time) and compliance (e.g., avoiding deviations or reducing

risks). In particular, the objective of process mining is to gain concrete and actionable process insights from event data.

The course starts with an introduction to the basics of business process management and presents the main drivers of process mining, which are the process models describing the behavior of a workflow and the event logs providing detailed information about the history of process executions.

The first part of the course covers the three main types of process mining.

(1) Process discovery techniques take an event log and produce a process model explaining the behavior recorded in the log without using any a-priori information. Various advanced process discovery techniques will be discussed to compare their strengths and weaknesses.

(2) Conformance checking checks if the behavior allowed by a process model conforms with the reality as recorded in an event log of the same process. Metrics and dimensions for conformance checking (fitness, precision, generalization, and simplicity) will be discussed, together with techniques to measure the alignment between process models and event logs.

(3) Enhancement aims at extending, refining or improving an existing process model using information about the actual process recorded in some event log. The course will investigate several techniques for process enhancement, such as techniques for the time prediction of running cases (bottleneck analysis) and for decision mining.

During the course, an academic software prototype (ProM) and an industrial one (Disco) will be used to test all the discussed process mining techniques against real-life and synthetic data sets. Besides learning theoretical concepts, participants will be exposed to event data from a variety of domains, including hospitals, insurance companies, etc.

The second part of the course will focus on the most recent scientific advances in process mining. In particular, different techniques for the discovery of BPMN models will be presented, as well as process discovery and conformance checking techniques based on declarative specifications. The latter are very suitable to be used to compactly represent business processes in exception-prone environments where several different execution paths are allowed.

Then, we will discuss how recent techniques from academic artificial intelligence research can be employed to advance the state-of-the art in process mining, specifically for aligning event logs against imperative and declarative process models. These approaches are able to outperform (even of several orders of magnitude) the existing ad-hoc approaches implemented in ProM.

# Learning Objectives, Prerequisites and Evaluation

The learning objectives of this course are:

- master the theoretical foundations underlying existing process mining techniques;
- understand how to practically apply basic process discovery / conformance checking / enhancement techniques, both manually and using tools;
- work with the process mining tools ProM and Disco to perform process mining analyses and correctly interpret the results;
- know the most recent scientific advancements in process mining.

This course is aimed at PhD students, Master students and professionals. A basic understanding of logic, set theory, and statistics (at the undergraduate level) is assumed. Basic computer skills are required to use the software provided by the course (but no programming experience is needed).

In order to get the credits provided by the course, a short project on process mining from the course participants (to be held individually or in group) is required. With the project, the participants must demonstrate the ability to put into practice the activities illustrated or carried out during the course.

# Lectures

## *Lecture 1 (4 hours) (Marrella)*

- **When:** 11 Feb 2019, 10:00-14:00, Room B203
- **Description**: In this lecture, we introduce the basics of business process management and the key features of process mining. Specifically, we discuss process models and event logs as means to understand and analyse the dynamic behaviour of business processes.
- **Topics**: Introduction to Business Process Management and Process Mining. Process Modeling: imperative and declarative models. Basics of Petri Nets, BPMN (Business Process Modeling and Notation) and Declare. Classroom exercises.

## *Lecture 2 (3 hours) (Maggi)*

- **When:** 12 Feb 2019, 10:00-13:00, Room B203
- **Description**: In this lecture, we introduce the basic technique to perform process discovery, named the α-algorithm. Then, we analyze alternative discovery approaches that allow users to tackle the limitations of the α-algorithm. Any presented discovery approach will be tested with the academic process mining tool ProM.

- **Topics**: Introduction to Process Discovery. Description of the α-algorithm to infer Petri nets from event logs. Discussion of alternative approaches to process discovery. Testing of process discovery approaches through ProM.

## *Lecture 3 (3 hours) (Marrella)*

- **When:** 13 Feb 2019, 10:00-13:00, Room B203
- **Description**: In this lecture we introduce the basic techniques, metrics and dimensions to check and measure the conformance of event logs against their underlying process model. Any presented conformance checking approach will be tested with the academic process mining tool ProM.
- **Topics**: Introduction to Conformance Checking. Metrics and dimensions for conformance checking (fitness, precision, generalization, and simplicity). Presentation of a technique for aligning event logs and process models. Testing of conformance checking techniques through ProM.

## *Lecture 4 (3 hours) (Maggi)*

- **When:** 14 Feb 2019, 10:00-13:00, Room B203
- **Description**: This lecture presents the most recent advances in the process mining field. In particular, process discovery with BPMN will be presented and tested with ProM. In addition, a wide range of process mining techniques with declarative specifications will be presented. These techniques will also be tested using the plug-ins of the process mining tool ProM.
- **Topics**: Introduction to Process Discovery and Conformance Checking with declarative specifications. Presentation of Process Discovery with BPMN models and introduction to more recent process discovery techniques like the Inductive Miner.

## *Lecture 5 (3 hours) (Marrella)*

- **When:** 15 Feb 2019, 10:00-13:00, Room B203
- **Description**: In this lecture we discuss how recent techniques from academic artificial intelligence research can be employed to advance the state-of-the art in conformance checking by outperforming the existing ad-hoc approaches implemented in ProM.
- **Topics**: Presentation of two recent techniques based on Automated Planning in AI for aligning event logs against imperative and declarative process models.

*Lecture 6 (4 hours) (Maggi)*

- **When:** 18 Feb 2019, 10:00-13:00, Room B203
- **Description**: This lecture covers additional process mining techniques such as performance analysis, social network analysis and process enhancement. In addition, the use of the industrial process mining tool Disco will be introduced.
- **Topics**: Introduction to performance analysis, social network analysis and process enhancement. Testing of process discovery approaches through Disco.

# References


- van der Aalst, Wil: “Process Mining. Data Science in Action”. Springer Berlin Heidelberg (2016).
- Wil M. P. van der Aalst et al.: “Process Mining Manifesto”. *Business Process Management Workshops* (1), pp. 169-194 (2011)
- van der Aalst, Wil, Adriansyah, Arya and van Dongen, Boudewijn: “Replaying history on process models for conformance checking and performance analysis”. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2.2*, pp. 182-192 (2012)
- De Leoni, M., Marrella, A.: “Aligning Real Process Executions and Prescriptive Process Models through Automated Planning”. *Expert Systems with Applications*, Volume 82, pp. 162–183, Elsevier (2017)
- De Giacomo, G., Maggi, F. M., Marrella, A, Patrizi, F.: “On the Disruptive Effectiveness of Automated Planning for LTLf-based Trace Alignment”. *AAAI 2017*, Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, USA, February 4-9 (2017).
- De Leoni, M., Maggi, F. M., van der Aalst Wil M. P.: “Aligning Event Logs and Declarative Process Models for Conformance Checking”. *BPM 2012*, pp. 82-97 (2012)
- Fabrizio Maria Maggi, R. P. Jagadeesh Chandra Bose, Wil M. P. van der Aalst: “Efficient Discovery of Understandable Declarative Process Models from Event Logs”. *CAiSE 2012*, pp. 270-285 (2012)
- Burattin A., Maggi, F. M., Sperduti A.: “Conformance checking based on multi-perspective declarative process models”. *Expert Systems with Applications*, Volume 65, pp. 194-211 (2016)
- Sander J. J. Leemans, Dirk Fahland, Wil M. P. van der Aalst: Discovering Block-Structured Process Models from Incomplete Event Logs. *Petri Nets 2014*, pp. 91-110 (2014)
- Conforti R., Dumas M., García-Bañuelos L., La Rosa M.: “BPMN Miner: Automated discovery of BPMN process models with hierarchical structure”. *Inf. Syst.*, Volume 56, pp. 284-303 (2016)

# Short Bio

**Fabrizio Maria Maggi** received his PhD degree in Computer Science in 2010, and after a period at the Architecture of Information Systems (AIS) research group - Department of Mathematics and Computer Science - Eindhoven University of Technology, he is currently a Senior Researcher at the Software Engineering Group - Institute of Computer Science - University of Tartu. His PhD dissertation was entitled "Process Modelling, Implementation and Improvement" and his areas of interest have included in the last years business process management, service-oriented computing, and software engineering. He authored more than 100 articles on process mining, (declarative) business process modeling and business constraints/rules, monitoring of business constraints at runtime, service oriented architectures, service choreographies and service composition. He was awarded with the best paper award of the BPM conference (the most prestigious conference in the field of Business Process Management) in 2015 and in 2016. He serves as senior program committee member of the same conference. In 2015, he was awarded with the best researcher award granted by the department of Computer Science of University of Tartu.

**Andrea Marrella** is currently research fellow at Sapienza Università di Roma, and recently got a national habilitation to associate professorship. His research activity focuses on how to integrate Artificial Intelligence with Business Process Management solutions, to untangle complex challenges such as the automated synthesis of process models, the automated adaptation of running processes and the optimal alignment of execution traces against their underlying (procedural or declarative) process models; such topics are challenged in the application domains of smart manufacturing, healthcare, emergency management and cybersecurity. In 2017, he received the best paper award at the 29th International Conference on Advanced Information Systems Engineering (CAiSE 2017). He is acting as Information Director of ACM Journal on Data and Information Quality. He is Principal Investigator of two research projects funded by Sapienza Università di Roma, and he collaborated to several EU-funded research projects, including WORKPAD (FP6) and Smart Vortex (FP7). His current h-index is 18.