Exercises 04: PM

Kevin Heubacher (15329079), Akshay Chikhalkar (15489036),
John Paul Semakula (15487087)

10.11.2022

1 Exercise 1: Compare Mining Algorithms

In the first exercise we are given three process mining algorithms and four metrics to analyse the performance of those algorithms. The given metrics are:
Replay fitness quantifies how well a process model is able to reproduce the behaviour that occurs in the event logs.
Precision quantifies how much of the behaviour in the process model is not seen in the event logs. Compares at certain points the set of transitions available in the process model with the activities found inside the event log.
Generalization assesses the extent to which the resulting process model will be able to reproduce future behavior of the process.
Simplicity quantifies the complexity of a process model. It considers the quantity of input and output arcs of each transition. The higher the amount of arcs the lower the simplicity.
The resulting value of each metric lies between $[0, 1] \in \mathbb{R}$. The higher the result the better. The analysed process mining algorithms are the Alpha Miner, Heurist Miner and Inductive Miner. Their performance is analysed by applying them on three different event logs. For each event log a table was created, documenting the performance results of the three process miners.

coffee.csv	Alpha Miner	Heurist Miner	Inductive Miner
Replay Fitness (T)	0.88	1.0	1.0
Replay Fitness (A)	0.93	1.0	1.0
Precision (T)	1.0	0.96	0.96
Precision (A)	1.0	0.96	0.96
Generalization	0.35	0.31	0.31
Simplicity	1.0	0.9	0.9

Table 1: Evaluation results for event log coffee.csv

The Heurist Miner and Inductive Miner performed identically for every metric on the coffee.csv event log. The Alpha Miner has a lower score in Replay Fitness, but performed better in Precision, Generalization and Simplicity.

roadtraffic50traces.xes	Alpha Miner	Heurist Miner	Inductive Miner
Replay Fitness (T)	0.81	0.92	1.0
Replay Fitness (A)	-	0.83	1.0
Precision (T)	0.65	0.88	0.58
Precision (A)	-	0.88	0.58
Generalization	0.62	0.55	0.65
Simplicity	0.8	0.67	0.68

Table 2: Evaluation results for event log roadtraffic50traces.xes

Alignment based Replay Fitness and Precision for the Alpha Miner were unable to be calculated on the roadtraffic50traces.xes event log. The Inductive Miner performed the best on Replay Fitness and Generalization. The Heurist Miner performed almost as good in Replay Fitness, but outperformed the other miners in Precision. Comparatively, the Alpha Miner only performed well in Simplicity.

roadtraffic100traces.xes	Alpha Miner	Heurist Miner	Inductive Miner
Replay Fitness (T)	0.80	0.93	1.0
Replay Fitness (A)	0.48	0.81	1.0
Precision (T)	0.82	0.89	0.43
Precision (A)	0.82	0.89	0.43
Generalization	0.52	0.53	0.65
Simplicity	0.91	0.62	0.68

Table 3: Evaluation results for event log roadtraffic100traces.xes

For the roadtraffic100traces.xes event log the Inductive Miner again performed best in Replay Fitness and Generalization. The Alpha and Heurist Miner performed much better in Precision.

Average performance	Alpha Miner	Heurist Miner	Inductive Miner
Replay Fitness (T)	0.83	0.95	1.0
Replay Fitness (A)	0.71	0.88	1.0
Precision (T)	0.82	0.91	0.66
Precision (A)	0.91	0.91	0.66
Generalization	0.50	0.46	0.54
Simplicity	0.90	0.73	0.75

Table 4: Average evaluation results

All the process miners performed poorly in Generalization. The Inductive miner is a clear winner in Replay Fitness, but is not as precise as the other miners. The Heurist and Alpha Miner performed well in every metric except Generalization.