Type: Master Thesis

Title: Place Net Oracle for Local Process Model Discovery

Description
As the popularity of process mining grew, so did the spectrum of processes that were analyzed. As a result, the traditional process discovery techniques were confronted with heavily unstructured processes they could not handle appropriately, and local process model discovery appeared. In contrast to process discovery, where the goal is to find one process model that explains all traces in the event log from start to end, local process model discovery aims to find a set of smaller models that describe what happens locally (describe subsequences of the traces that are of some maximal length) in the event log. Although the primary purpose of local process models was to give insights on event logs for which the traditional process discovery techniques failed to return a well-structured model, with time, the importance and application of local process models grew and became multi-fold. Now they are used for event abstraction, trace clustering, outcome prediction, etc.

There exist multiple approaches for discovering local process models. The approach significant for this thesis builds local process models by combining place nets\(^1\). The place nets are accepted as input of the algorithm and are generated by an oracle. The built local process models can be filtered and sorted based on different attributes or evaluation metrics. Currently, as a place oracle can be used any process discovery algorithm. However, there are multiple disadvantages to using traditional process discovery algorithms as place oracles. First, they do not consider locality. Second, their goal is to return a model and not a set of all possible place nets. The most usable place oracle currently is the eST miner. However, like all others, it does not consider locality, and time-wise does not scale well on large event logs.

The goal of this master thesis is to generate place nets given some event log as input. The master thesis includes investigating and experimenting with different heuristics for place net generation. We expect a ProM plugin (https://www.promtools.org) as an outcome of the master thesis. The student should also evaluate the new generation approaches on real and artificial event logs and compare them to existing methods and a random generator. Those that want the challenge can also formally reason why some heuristics are better than others.

---
\(^1\) Place nets are Petri nets with only one place connected to multiple transitions.
Prerequisites:
The thesis includes implementation as well as formal reasoning. Java programming skills and an interest in theoretical and practical aspects of process mining is required.

Pointers:
- Process Mining Book
- Coursera Process Mining Course
- Finding Complex Process-Structures by Exploiting the Token-Game
- From Place Nets to Local Process Models
- ProM

Supervisor:
Prof.dr.ir. Wil van der Aalst

Advisor:
Viki Peeva

For more Information:
Send an e-mail to peeva@pads.rwth-aachen.de. Make sure to include detailed information about your background and scores for completed courses.

Chair of Process and Data Science
Ahornstr. 55 (Eingang Mies-van-der-Rohe-Str.), Erweiterungsbau E2
52074 Aachen
Phone: +49 241 80 21 901
http://www.pads.rwth-aachen.de