VU Business Intelligence 2

Faculty of Computer Science, University of Vienna

Univ. Prof. Dr. Maria Leitner
Dr. Erion Çano

04.11.2022

Exercise 2: Alpha Algorithm Implementation (16 Points)

Goal: Implement the alpha algorithm in teams of 2 people.

Therefore, we will publicly available data sets:

- The Business Process Intelligence (BPI) Challenge 2012
    - http://dx.doi.org/10.4121/uuid:3926db30-f712-4394-aebc-75976070e91f
    - BPIC 2012: https://www.win.tue.nl/bpi/doku.php?id=2012:challenge
- The Business Process Intelligence (BPI) Challenge 2017
    - https://data.4tu.nl/articles/dataset/BPI_Challenge_2017/12696884

You may also use the example log files uploaded in Moodle (recommendation: only nonoise).

Approach

Please implement in a team of 2 the alpha algorithm (see references below). You may use libraries for the data structures (e.g., PM4PY or others) but you may not use them for applying the alpha algorithm. The alpha algorithm should be independently programmed by yourself. You may use libraries to compare your results in the process.

Also calculate the Replay Fitness for each of your resulting process models. Specifically, the Average Trace Fitness und Log fitness. Please compare it with other process mining algorithms provided by the library, specifically: the alpha, alpha plus, heuristic and inductive miner. Add at least one sentence for the interpretation of the comparison in the report.

Document all your findings in a report:

- Describe all relevant results and your approach. It is recommended to use a similar structure:
    - Title: Alpha algorithm implementation
    - Group number, Student names and IDs,
    - Approach
        - Programming language, libraries used
        - Use of data structures
        - Challenges (if applicable)
        - How to run the program
    - Results
        - using the BPI 2012 and BPI 2017 data sets. Include at least a figure from your resulting process models (BPI 2012 and BPI 2017).
        - Export these visualizations as Petri net (PNML file), See Submission below.
    - Team work
        - Describe how you worked together and divided the tasks

Delivery Talks (Abgabegespräche)

To receive points in this exercise, this exercise requires a delivery talk (in presence). Registration of the delivery talks is via Moodle and are expected to take place in December 2022. Please expect to present in max. 7 minutes your approach and your results. You may use the uploaded report or another presentation deck for this meeting.

Distribution of points

Given that the algorithm has been implemented and is working correctly and the delivery talk has been participated, the points are allocated as follows:

- Correct working alpha algorithm (11)
- Clear documentation in report and presentation (5)

Important notes:

- All source code will be investigated for plagiarism. Please be careful and work in your own teams.
- Please also name the tool you used for your investigation.
- Do not include the BPI challenge data sets in your exercise submissions!
- Please include page numbers in your report.

References

alpha algorithm: de Medeiros, A. K. A., van der Aalst, W. M., & Weijters, A. J. M. M. (2003, November). Workflow mining: Current status and future directions. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"(pp. 389-406). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-39964-3_25

Submission and naming conventions:

- Report (PDF): `TeamNumber_Exercise_ExerciseNumber_Report.PDF`
- Source Code (ZIP) : `TeamNumber_Exercise_ExerciseNumber_Source.ZIP`
- Exported PNML files of the resulting Petri nets (ZIP) : `TeamNumber_Exercise_ExerciseNumber_PNML.ZIP`

Please upload everything in the stated file types.