

A Python Tool for Object-Centric Process Mining Comparison (Extended Abstract)

Anahita Farhang Ghahfarokhi
Process and Data Science (Informatik 9)
RWTH Aachen University
Aachen, Germany
farhang@pads.rwth-aachen.de

Wil M.P. van der Aalst
Process and Data Science (Informatik 9)
RWTH Aachen University
Aachen, Germany
wvdaalst@rwth-aachen.de

Abstract—Object-centric process mining provides a more holistic view of processes where we analyze processes with multiple case notions. However, most object-centric process mining techniques consider the whole event log rather than the comparison of existing behaviors in the log. In this paper, we introduce a stand-alone object-centric process cube tool built on the PM4PY-MDL process mining framework. Our infrastructure uses both object and event attributes to build the process cube which leads to different types of materialization. Furthermore, our tool is equipped with the state of the art object-centric process mining techniques. Through our tool the user can visualize the extracted object-centric event log from process cube operations, export the object-centric event log, discover the state-of-the-art object-centric process model for the extracted log, and compare the process models side-by-side.

Index Terms—Object-Centric Process Mining, Object-Centric Event Logs, Process Comparison, Process Cubes

I. INTRODUCTION

Process mining is a field of data science that aims to bridge the gap between business process model-based analysis and data-oriented analysis. Process mining techniques include process discovery, conformance checking, and process enhancement methods [10].

Event logs are the starting point to apply process mining techniques. Event logs consist of events where each event refers to one case notion, activity, timestamp, and some additional attributes such as resource. Common process mining techniques are based on event logs with one case notion, however, in reality, several case notions are involved in one process, e.g., a simple Order-to-Cash (O2C) process where orders, offers, and invoices are involved. Object-centric process mining is a novel branch of process mining that aims to develop process mining techniques on top of Object-Centric Event Logs (OCELs) [6]. Initial approaches have been developed to extract OCEL logs from information systems [1], [4], [9] and discover process models from OCEL logs [3], [7]. However, there may exist a variety in object-centric processes that requires the separation of different processes from each other [5]. Therefore, process cubes are introduced that are inspired by the notion of OLAP and are developed to compare processes with each other through process cube operations such as slice and dice [12]. Several implementations of process cubes are developed [2], [8]. However, they cannot support event logs with multiple objects, i.e., OCEL logs.

In this demo paper, we present an interactive tool that permits the user to compare object-centric processes with each other through process cube operations. Furthermore, it permits to discover object-centric process models (object-centric Directly Follows Graphs [1], Object-Centric Petri Nets [11]) and compare the process models side-by-side.

The remainder of this paper is organized as follows. In Section II, we describe the main functionalities of the tool that are provided. In Section III, we evaluate the scalability of the tool. Finally, Section IV concludes the paper and presents some future work for the extension of our work.

II. OBJECT-CENTRIC PROCESS CUBES

In our open-source tool¹, called OCPC (Object-Centric Process Cube), we have used Tk GUI toolkit as the user interface which is one of the popular standard Python interfaces. A snapshot of the tool is shown in Figure 1 where three sub windows are highlighted:

- **Input:** the input is JSON-OCEL/XML-OCEL. In OCEL, an event is related to event attributes and some possible objects related to that event. Furthermore, each object can have its own properties in another table. Therefore, as shown in Figure 1, in the first highlighted sub window, we have shown events with event attributes and their objects in one table, whereas, objects within their attributes are shown in separate tables. The user can select event attributes and object attributes as the dimensions of the process cube and build the cube.
- **Wizard:** After creating the cube, we see overviews of the cube in the second sub window. It is possible to select dimensions of the cube as rows and columns and see the number of events in the selected slices/dices. The combo box "Materialization" is related to design choices that we have in this tool for materializing that did not exist in the previous process cubes. In the "Existence" option for materialization, there should exist an object in that event that satisfies the property. For example, in the left-view in Figure 1, the option is "Existence" and the highlighted green dice in this view shows in 557 events *Echo show 8* was involved. In the "All" option

¹<https://github.com/AnahitaFarhang/object-centric-process-cube>

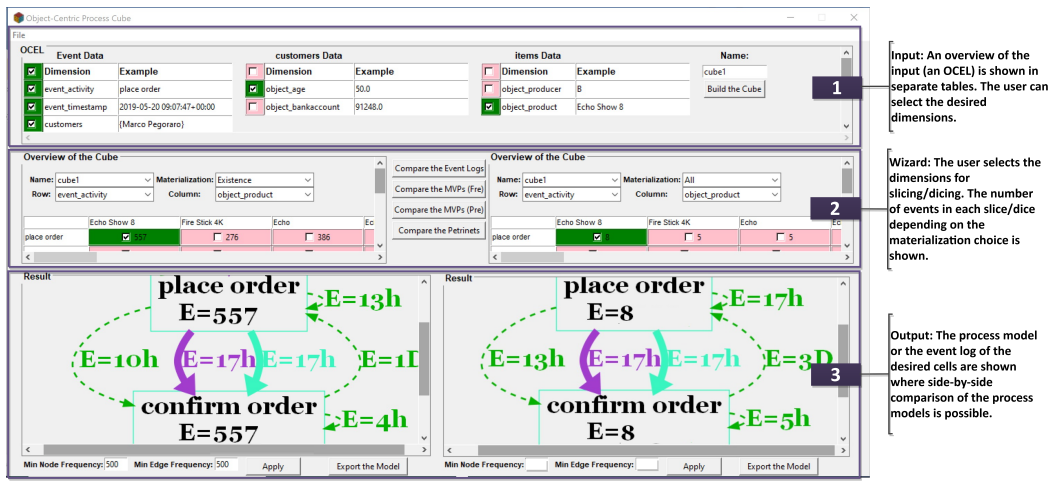


Fig. 1. The user interface of the tool that is developed in Tkinter GUI.

TABLE I
SCALABILITY OF THE TOOL FOR DIFFERENT NUMBERS OF EVENTS:
N_EVENT ATTRIBUTES=4, AND N_OBJECT ATTRIBUTES=4

Number of events	1000	5000	10000	15000	20000
Time of building the cube (s)	11.15	120.23	237.26	361.44	490.12

TABLE II
SCALABILITY OF THE TOOL FOR DIFFERENT NUMBERS OF EVENT
ATTRIBUTES: N_EVENTS=20000, AND N_OBJECT ATTRIBUTES=4

Number of object attributes	0	1	2	3
Time of building the cube (s)	60.17	127.27	254.79	490.12

for materialization, all the objects, in the event, should satisfy that property. An example is shown in right-view where this option is "All". The highlighted green cell in this view shows that in 8 events all the items were *Echo show 8*.

- Output: In the third sub window, we have compared the process models of the selected cells shown in the second sub window. As we see, there are differences in the duration of activities for the selected cells. Furthermore, depending on the operation selection in the second sub window, it is possible to compare the extracted event logs, and Object-Centric Petri nets of the selected cells.

III. SCALABILITY

The scalability of the tool in terms of number of events, event attributes, and object attributes are shown in Tables I, II, and III, respectively². The result of the analysis with different settings shows the time required for creating the cube increases linearly, linearly, and non-linearly when increasing the number of events, event attributes, and object attributes, respectively. These relationships are justifiable by the nature of OCELS where an event can contain multiple objects and there is a one-to-many relationship between an event and its objects which results in a non-linear relationship shown in Table III.

IV. CONCLUSION

Here, we present an interactive object-centric process cube tool that enables the exploration, discovery of object-centric

²All the above experiments were performed on a laptop with the specifications: PC Intel(R) Core(TM) i7-8550U CPU @ 1.80GHz.

TABLE III
SCALABILITY OF THE TOOL FOR DIFFERENT NUMBERS OF OBJECT
ATTRIBUTES: N_EVENTS=20000, AND N_EVENT ATTRIBUTES=4

Number of object attributes	0	1	2	3	4
Time of building the cube (s)	62.16	99.43	233.63	270.64	490.12

process models along with side-by-side model comparison. Using the tool that is implemented in PM4PY-MDL, users explore object-centric event logs using process cube operations, export the partitioned OCEL, and discover the object-centric process models, annotated with frequency and performance, which helps in understanding processes better. A video displaying the functionalities of our tool is available at the address <https://youtu.be/zenHt3wdZP4>. We aim in the future to extend this tool with an automatic method to choose interesting slices and dices to explore the cube.

REFERENCES

- [1] A. Berti and W.M.P. van der Aalst. Extracting multiple viewpoint models from relational databases. In *SIMPDA*, volume 379, pages 24–51, 2019.
- [2] A. Bolt and W.M.P. van der Aalst. Multidimensional process mining using process cubes. In *BPMDS*, volume 214, pages 102–116, 2015.
- [3] D. Cohn and R. Hull. Business artifacts: A data-centric approach to modeling business operations and processes. *IEEE Data Eng. Bull.*, 32:47–59, 2009.
- [4] E.G.L de Murillas. Process mining on databases: Extracting event data from real-life data sources. 2019.
- [5] A. Farhang, A. Berti, and W.M.P. van der Aalst. Process comparison using object-centric process cubes. *arXiv:2103.07184*, 2021.
- [6] A. Farhang, G. Park, A. Berti, and W.M.P. van der Aalst. OCEL: A standard for object-centric event logs. In *SIMPDA*, pages 169–175. Springer, 2021.
- [7] E.H.J. Nooijen, B.F. van Dongen, and D. Fahland. Automatic discovery of data-centric and artifact-centric processes. In *BPM Workshops*, volume 132, pages 316–327, 2012.
- [8] J.T.S. Ribeiro and A.J.M.M. Weijters. Event cube: Another perspective on business processes. In *OTM*, volume 7044, pages 274–283, 2011.
- [9] A.P. Simović, S. Babarogić, and O. Pantelić. A domain-specific language for supporting event log extraction from ERP systems. In *ICCCC*, 2018.
- [10] M.S. Uysal, S.J. van Zelst, T. Brockhoff, A. Farhang, M. Pourbafrani, R. Schumacher, S. Junglas, G. Schuh, and W.M.P. van der Aalst. Process mining for production processes in the automotive industry. In *Industry Forum at BPM*, 2020.
- [11] W.M.P. van der Aalst and A. Berti. Discovering object-centric petri nets. *Fundamenta Informaticae*, 175:1–40, 2020.
- [12] T. Vogelgesang and H.J. Appelrath. Multidimensional process mining: A flexible analysis approach for health services research. In *EDBT/ICDT*, pages 17–22, 2013.