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

Anahita Farhang Ghahtfarokhi  
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], [3], [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 tool[1] 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

https://github.com/AnahitaFarhang/object-centric-process-cube
The scalability of the tool for different numbers of events: $N_{\text{events}}=4$, and $N_{\text{object attributes}}=4$.

<table>
<thead>
<tr>
<th>Number of events</th>
<th>1000</th>
<th>5000</th>
<th>10000</th>
<th>15000</th>
<th>20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time of building the cube (s)</td>
<td>11.15</td>
<td>120.12</td>
<td>257.26</td>
<td>401.44</td>
<td>490.12</td>
</tr>
</tbody>
</table>

The scalability of the tool for different numbers of event attributes: $N_{\text{events}}=20000$, and $N_{\text{object attributes}}=4$.

<table>
<thead>
<tr>
<th>Number of object attributes</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time of building the cube (s)</td>
<td>62.16</td>
<td>99.43</td>
<td>253.63</td>
<td>230.64</td>
<td>490.12</td>
</tr>
</tbody>
</table>

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](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


---

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