

### Question-1:

Open the event log ('Receipt phase of an environmental permit application process (\_WABO\_) CoSeLoG project.fbt') in Disco and switch to the 'Statistics' view.

Without switching to other views, use the statistics view to answer the following three subquestions:

1. How many events are there on average per case?
2. Can you indicate whether each case seems to be unique or whether many cases follow the same activity sequence?
3. What is the main observation that can be made from the 'Events over time' graph?

### Question-2

While still in Disco, switch to the 'map' view to display a process map.

Using the map view, change the activity and path detail settings in order to create a comprehensible process map (e.g. a process map that could be printed on one A4 or letter paper or shown on a single computer screen while still being readable in full).

1. Discuss this process map, what is the main process?
2. Which activities and paths between activities are frequent?

In your answer, include the settings you used for both the activity and path sliders.

### Question-3

While still in Disco, and while using the same process map (e.g. do not change the activity and path settings), switch to the performance projection.

Discuss where the process takes most time, e.g. where there are possibilities for improvement. Relate these times (of the bottlenecks) to the time spent in other parts of the process. In other words, discuss how severe the bottleneck is with respect to the time spent on other activities.

Also explicitly mention the performance metric chosen (e.g. total, mean, median, or max) and why you have chosen this setting.

### Question-4

Now load the original event log in ProM. Visualize the event log using the Dotted Chart or XDottedChart visualizer (by pressing the 'eye'-icon with the event log selected and switching to the Dotted Chart or XDottedChart visualizer).

Using the Dotted Chart, answer the following questions:

1. Is the arrival rate of new cases constant? If not, when are there fluctuations? If yes, how can we see this from the Dotted Chart?
2. Can you observe a change in the global process?

Note that you don't need to change the component, time or coloring settings. You can however re-sort the traces on the time of the first event, and zoom in or out if you want.

The Dotted Chart is explained in lecture 4.8: 'Exploring Event Data'.

### Question-5

You are now asked to discover a Petri net on the event log. However, the unfiltered event log results in an incomprehensible Petri net. Therefore, you are allowed to run the 'Filter log using simple heuristics' plug-in *once* on the original event log to discover a Petri net on the filtered event log.

1. Clearly indicate which settings you have used for the 'Filter log using simple heuristics' plug-in.
2. Explicitly motivate the filtering settings chosen, why did you pick this percentage or selection of activities?
3. Discuss and argue which plug-in (or chain of plug-ins) you have used to discover a Petri net, for instance by comparing two or more plug-in results and arguing why one of the Petri nets is better.
4. Explain the (best) Petri net: what is the main process and what are notable parts of the Petri net?

Note that this question requires you to experiment with different filtering settings and discovery plug-ins. You are not required to describe *everything* you have tried but found unsuccessful. Only describe the successful combination of plug-ins and its result(s) and argue why your final result is 'good'.

#### **Suggested list of plug-ins or plug-in chains to produce a Petri net:**

- Mine for a Petri Net using Alpha-algorithm
- Mine for a Petri Net using ILP
- Mine for a Heuristics Net using Heuristics Miner *followed by* Convert Heuristics net into Petri net
- Mine for a Petri net with Inductive Miner

### Question-6

The organization has a process model that describes the 'should be' process (i.e. a normative process model). Load the file 'normativeModel.pnml' into ProM and apply conformance checking on this process model, and on the full unfiltered original event log.

1. Include a screenshot of the part of the normative process model, with the conformance information projected onto it, that shows where most of the deviations occur.
2. What is the replay fitness (the 'trace fitness' statistic) of the event log on the normative process model?

3. Select the transition 'T06 Determine necessity of stop advice+complete' (on the top left of the model) and discuss its element statistics: how many times is the transition executed correctly and how many times incorrectly?
4. Using the element statistics of transition 'T06 Determine necessity of stop advice+complete', what can you say about the (in)correct execution of this activity?

**Instructions to align the process model with the event log:**

1. Import the normative model using the 'PNML Petri net files' importer.
2. Select the imported normative Petri net and the event log, start the plug-in called 'Replay a Log on Petri Net for Conformance Analysis' (not the variant with performance!), and click 'yes' in the 'No Final Marking' pop-up.
3. Select the 'sink' place on the left (note: do not select '0-sink' etc.) and click the button 'Add Place >>' to add the place 'sink' to the candidate final marking list. Now click 'Finish'.
4. Click 'Finish' in the mapping wizard.
5. Click 'No, I've mapped all necessary event classes' to indicate that some events are not present in the normative model.
6. Now click 'Next' and 'Finish'. The normative process model is shown with conformance information projected onto it.

If you followed these instructions exactly you do not need to mention these steps in your answer.

More information regarding this conformance technique is provided in lecture 4.7: 'Aligning observed and modeled behavior' (and to a lesser extend in the lectures 4.3 through 4.6).

**Question-7**

The final analysis you have to perform on the original event log is a resource analysis, e.g. looking at the user behavior in the event log.

1. Use the plug-in 'Mine for a Subcontracting Social Network'. Note that subcontracting means that if individual  $j$  frequently executed an activity in-between two activities executed by individual  $i$ , then individual  $i$  subcontracted work to individual  $j$ . Answer the following question using this view: Can two or more groups of users be distinguished? Explicitly discuss the settings you have used in the resulting visualization.
2. Again use one of the two Dotted Chart plug-ins. For the XDottedChart change the component type to 'org:resource'. If you use the Dotted Chart visualizer change the 'Y Axis Attribute' to 'C: Resource classifier' and the color attribute to 'C: Activity Classifier'. Answer the following two questions using this view:
3. Are all users executing activities from the start of the event log, or are some users joining later?
4. Are users mainly executing particular activities or are most users executing most of the activities?

**Question-8**

To conclude this assignment, briefly discuss three observations you have made during your analysis that you would like to communicate to the business user.

Think for instance of possible improvement opportunities and starting points for further investigation.