

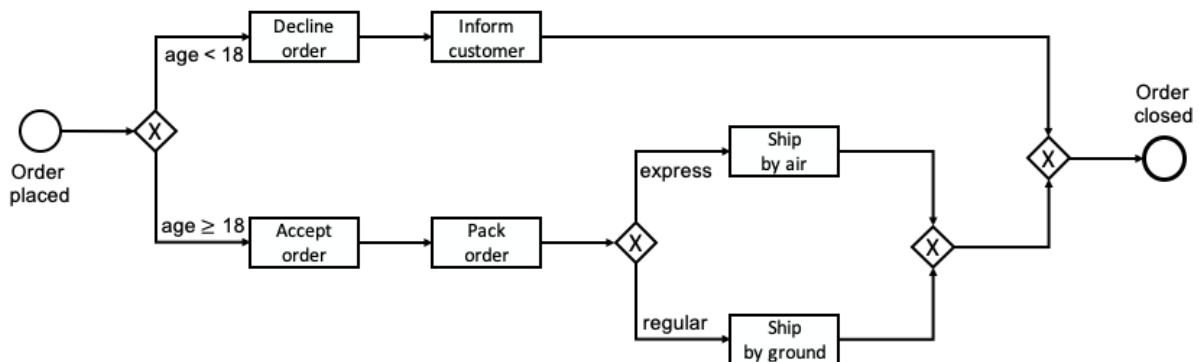
Advanced Process Mining

Summer term 2020

Exercise sheet 10

Multi-Perspective Alignments • Simplicity

Exercise 1: Multi-Perspective Alignments



Given the BPMN diagram above, calculate the deviation cost for the traces in the following event log using *multi-perspective alignments*:

Case ID	Timestamp	Activity	Age	Express
1	01.06.20 - 18:00:42	order placed	16	true
2	01.06.20 - 18:00:50	order placed	32	false
1	01.06.20 - 18:01:02	decline order	16	true
1	01.06.20 - 18:01:12	inform customer	16	true
1	01.06.20 - 18:01:12	order closed	16	true
2	01.06.20 - 18:02:20	accept order	32	false
3	01.06.20 - 22:01:27	order placed	15	true
3	01.06.20 - 22:33:57	accept order	15	true
2	02.06.20 - 08:43:44	pack order	32	false
2	02.06.20 - 13:22:53	Ship by air	32	false
3	02.06.20 - 13:41:02	Ship by ground	15	true
2	02.06.20 - 13:50:00	Order closed	32	false
3	02.06.20 - 13:50:01	Order closed	15	true

Solution

σ_1	order placed	decline order age: 16	inform customer	order closed	
N	order placed	decline order age: < 18	inform customer	order closed	
σ_2	order placed	accept order age: 32	pack order	ship by air express: false	order closed
N	order placed	accept order age: ≥ 18	pack order	ship by air express: true	order closed
σ_3	order placed	accept order age: 15	\gg	ship by ground express: true	order closed
N	order placed	accept order age: ≥ 18	pack order	ship by ground express: false	order closed

$$C_1 = 0$$

$$C_2 = 1$$

$$C_3 = 3$$

Exercise 2: True or False

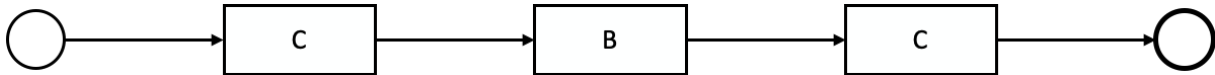
- For a successful process mining project it is not essential to have a common understanding of the objective.
- The terms *Multi-Perspective alignments* and *Multiple Trace alignments* describe the same thing and can be used synonymously.

Solution

- False.** Before a process mining project is started, a concrete use case and a clear purpose of the project has to be identified.
- False.** Compared to the simple alignment method the multi-perspective alignment takes additional attributes into account.
Multiple Trace alignments is not only optimising the alignment cost for a single trace but attempts to minimise the overall alignment cost.

Exercise 3: Simplicity

- What is quantified by simplicity and how is it measured?



- b) Transform the process model above into a process tree.
- c) Calculate the simplicity Q_S of the resulting process tree using the formula and method provided by *Buijs et al.* in Quality Dimensions in Process Discovery [...] 2014.

$$Q_S = 1 - \frac{\#duplicate\ activities + \#missing\ activities}{\#nodes\ in\ process\ tree + \#event\ class\ in\ event\ log}$$

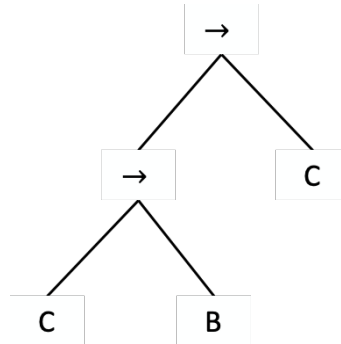
Event log:

Trace	#
ABCDEG	6
ABCDG	38
ABDCEG	12
ABDCFG	26
ABCDFG	8
ACBEG	1
ADBCFG	1
ADBCEG	1
ADCBFG	4
ACDBFG	2
ACBFG	1

- d) According to the approach off by *Buijs et al.*, how does a process tree look like with:
- excellent simplicity
 - very poor simplicity

Solution

- a) The complexity of the model is quantified by the simplicity and is measured by comparing the size of the tree with the number of activities in the log.
- b) The resulting process tree is the following:



- c) $S_Q = 1 - \frac{1+6}{3+7} = 0.4$
- d) i) To achieve excellent simplicity, the number of duplicate activities and the number of missing activities has to be equal to zero.
If the denominator is much greater than the numerator, a good simplicity can still be achieved, even if the number of duplicate activities or missing activities is not zero. Then either the number of nodes in the process tree has to be very high, or the event log has to consist of a high number of different activities
- ii) If most of the nodes in the process tree are duplicates or many activities are missing in the process tree, then the simplicity of the process tree will be very poor.