**Ludwig-Maximilians-Universität München**
**Institut für Informatik**
Prof. Dr. Thomas Seidl
Janina Sontheim, Maximilian Hünemörder

# Knowledge Discovery in Databases

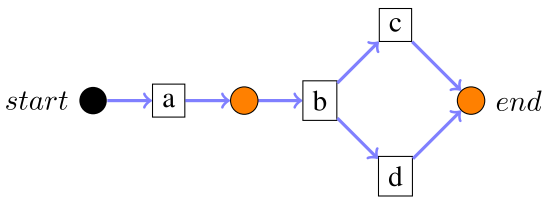
WS 2019/20

## Exercise 12: Process Mining

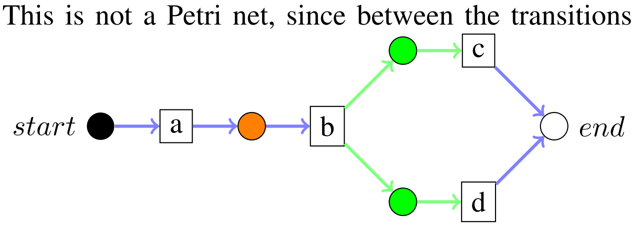
### Exercise 12-1 Petri Nets

Explain if the following graphs are petri nets, workflow nets, or even sound workflow nets.
Further express the graph as a sound workflow net if the graph is not yet a sound workflow net.

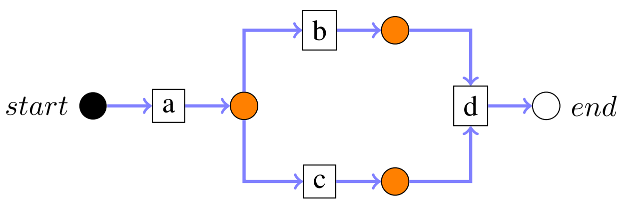
(a)



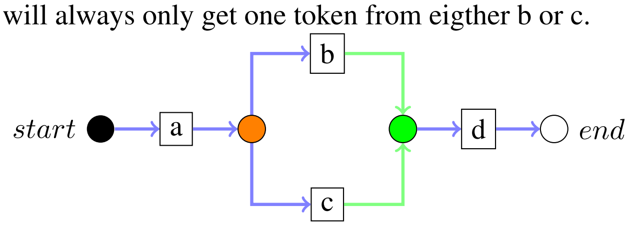

This is not a Petri net, since between the transitions b and c, resp. b and d there is a missing place.



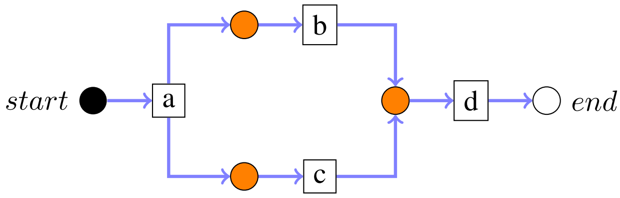

(b)



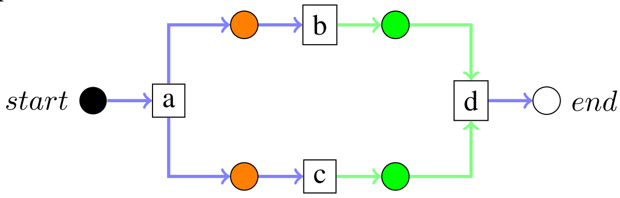

This is a Petri net and a workflow net, but it is not sound, since there are **deadlocks**. The OR-Split after the transition a gets closed by an AND-JOIN which causes a deadlock. Transition d can not fire since it will always only get one token from eigther b or c.




(c)




This is a Petri net and a Workflow net but it is not sound, due to the violation of the **proper completion** constraint. It can terminate while there is still a token in an other place than the output place.

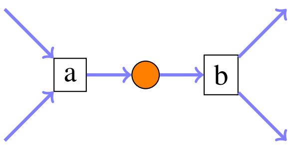



## Exercise 12-2 $\alpha$-Miner

(a) For the $\alpha$-Miner algorithm, we use the relations $>, \rightarrow, \|, \#$ to denote direct successions, causality, parallelism or choice. Considering the set of activities $\{a, b, c\}$, give notion (graphically) about the following patterns and associate the right relations with them according to the activities having been used:

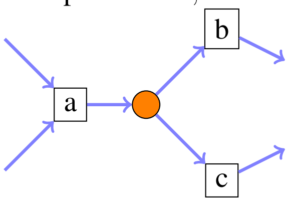
- Sequence Pattern

  Sequence: $a \rightarrow b$

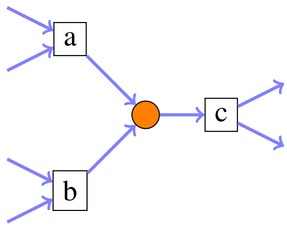
  
  

- XOR-Split and XOR-Join pattern

  XOR-split: $a \rightarrow b, a \rightarrow c$ and $b\#c$

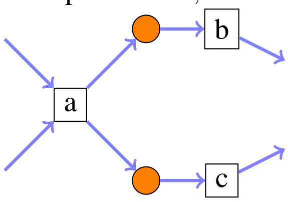
  
  

  XOR-join: $a \rightarrow c, b \rightarrow c$ and $b\#c$

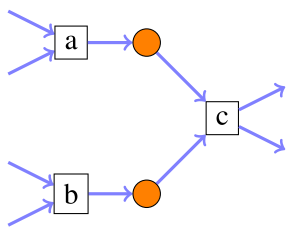
  
  

- AND-split and AND-join pattern

  AND-split: $a \rightarrow b, a \rightarrow c$ and $b\|c$

  AND-join: $a \rightarrow c, b \rightarrow c$ and $a\|b$

(b) Given the trace $L_1 = [\langle a, b, c, d\rangle, \langle a, c, b, d\rangle, \langle a, e, d\rangle]$. Determine the following sets:

- Set of activities: $T_L = \{t \in T | \exists_{\sigma \in L} t \in \sigma\}$

  Each activity in $L$ corresponds to a transition in $\alpha(L)$: $T_L = \{a, b, c, d, e\}$

- Set of start activities: $T_I = \{t \in T | \exists_{\sigma \in L} t = first(\sigma)\}$
  Fix the set of start activities - that is, the first elements of each trace: $T_I = \{a\}$
- Set of end activities: $T_O = \{t \in T | \exists_{\sigma \in L} t = last(\sigma)\}$
  Fix the set of end activities - that is, elements that appear last at a trace: $T_O = \{d\}$
- Set of paired activities:

$$\begin{aligned} X_L = \{(A,B) | A \subseteq T_L \wedge A \neq \emptyset \wedge \\ B \subseteq T_L \wedge B \neq \emptyset \wedge \\ \forall_{a \in A} \forall_{b \in B} a \rightarrow_L b \wedge \\ \forall_{a_1, a_2 \in A} a_1 \#_l a_2 \wedge \forall_{b_1, b_2 \in B} b_1 \#_L b_2\} \end{aligned}$$

  Find pairs $(A, B)$ of sets of activities such that every element $a \in A$ and every element $b \in B$ are causally related (i.e. $a \rightarrow_L b$), all elements in $A$ are independent ($a_1 \# a_2$), and all elements in $B$ are independent ($b_1 \#_L b_2$) as well:

$$\begin{aligned} X_L = &\{(\{a\},\{b\}),(\{a\},\{c\}),(\{a\},\{e\}),(\{a\},\{b,e\}), \\ &(\{a\},\{c,e\}),(\{b\},\{d\}),(\{c\},\{d\}),(\{e\},\{d\}), \\ &(\{b,e\},\{d\}),(\{c,e\},\{d\})\} \end{aligned}$$

- Set of paired activities that are maximal:

$$Y_L = \{(A,B) \in X_L | \forall_{(A',B') \in X_L} A \subseteq A' \wedge B \subseteq B' \implies (A,B) = (A',B')\}$$

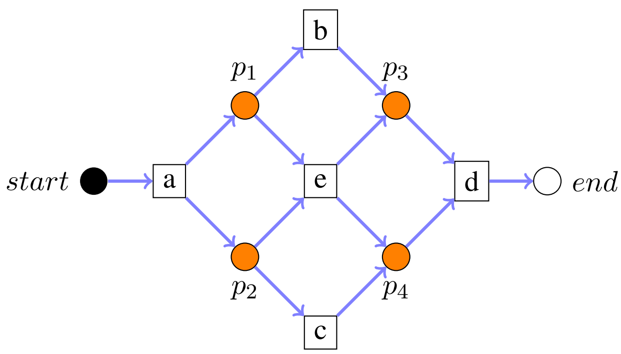
  Delete from $X_L$ all pairs $(A, B)$ that are not maximal.

$$Y_L = \{(\{a\},\{b,e\}),(\{a\},\{c,e\}),(\{b,e\},\{d\}),(\{c,e\},\{d\})\}$$

- Set of places: $P_L = \{p_{(A,B)} | (A,B) \in Y_L\} \cup \{i_L, o_L\}$
  Determine the place set: Each element $(A, B)$ of $Y_L$ is a place. To ensure the worklow structure, add a source place $i_L$ and a target place $o_L$. $p_L = \{p_{(\{a\},\{b,e\})}, p_{(\{a\},\{c,e\})}, p_{(\{b,e\},\{d\})}, p_{(\{c,e\},\{d\})}\}$
- Flow relations:

$$\begin{aligned} F_L = &\{(a, p_{(A,B)}) | (A,B) \in Y_L \wedge a \in A\} \cup \\ &\{(p_{(A,B)}, b) | (A,B) \in Y_L \wedge b \in B\} \cup \\ &\{(i_L, t) | t \in T_I\} \cup \{(t, o_L) | t \in T_O\} \end{aligned}$$

  Determine the flow relation: Connect each place $p_{(A,B)}$ with each element $a$ of its set $A$ or source transition and with each element of its set $B$ of target transitions. In addition, draw an arc from the source place $i_L$ to each start transittion $t \in T_I$ and an arc from each end transition $t \in T_O$ to the sink place $o_L$.



- Definition (*no task*): $\alpha$-Miner on event log $L$ is then defined as: $\alpha(L) = (P_L, T_L, F_L)$

(c) Construct the Footprint Table for trace $L_1$.

| | a | b | c | d | e |
|---|---|---|---|---|---|
| a | $\#_{L_1}$ | $\rightarrow_{L_1}$ | $\rightarrow_{L_1}$ | $\#_{L_1}$ | $\rightarrow_{L_1}$ |
| b | $\leftarrow_{L_1}$ | $\#_{L_1}$ | $\Vert_{L_1}$ | $\rightarrow_{L_1}$ | $\#_{L_1}$ |
| c | $\leftarrow_{L_1}$ | $\Vert_{L_1}$ | $\#_{L_1}$ | $\rightarrow_{L_1}$ | $\#_{L_1}$ |
| d | $\#_{L_1}$ | $\leftarrow_{L_1}$ | $\leftarrow_{L_1}$ | $\#_{L_1}$ | $\leftarrow_{L_1}$ |
| e | $\leftarrow_{L_1}$ | $\#_{L_1}$ | $\#_{L_1}$ | $\rightarrow_{L_1}$ | $\#_{L_1}$ |