Process Mining - 02269

Lecture 2

Limitations of Alpha Algorithm

Andrea Burattin

Log Completeness

- Required completeness for the α Algorithm
  - Completeness w.r.t. direct success relationship ($>_W$)
    - Whenever two tasks may occur as direct successors, this must be observed in at least one trace
  - Example: concurrent execution of five tasks, A, B, C, D, E
  - If $\langle A, B, C, D, E \rangle$ and $\langle B, A, C, E, D \rangle$ are observed, $\langle B, A, C, D, E \rangle$ is not required for log completeness!

- This completeness criterion reduces the number of required traces for highly concurrent processes dramatically

Rediscovery Problem

For which class of WF-nets can we guarantee that WF_1 and WF_2 are equivalent if logging is complete according to introduced notion?

Structural and behavioural assumptions on WF_1

What can be rediscovered?

The α Algorithm can rediscover

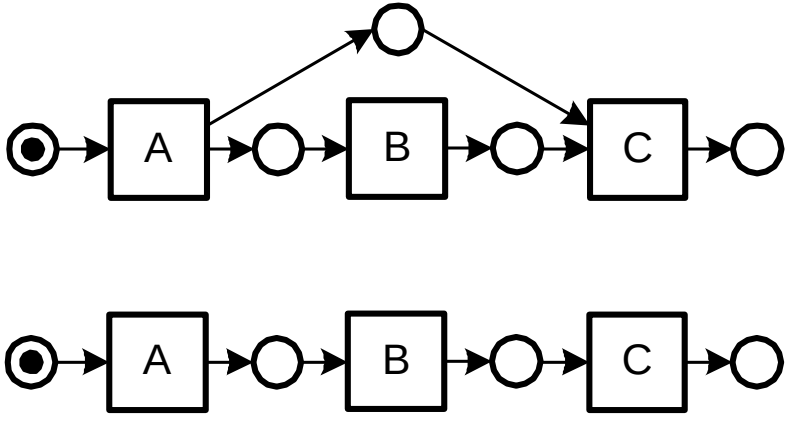
***sound structured workflow net without short loops**,*

if the event log is complete according to introduced notion

(based on direct success relationship)

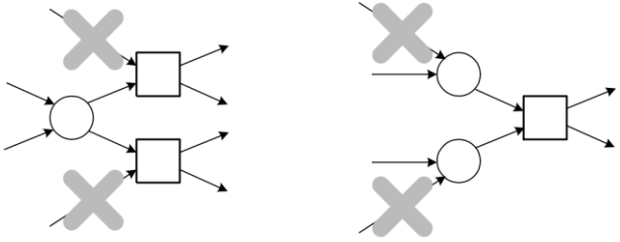
Implicit Places

- The presence or absence of implicit places does not change the behaviour of a net system
  - Hence, process models with implicit places cannot be re-discovered
  - Not really an issue, since there are no consequences for the behaviour



Structured Workflow-Nets

- Structured WF-nets (SWF-nets) are structural subclass:
  - A WF-net $N = (P, T, F)$ is a structured Workflow net iif
    - For all $p \in P$ and $t \in T$ with $(p, t) \in F$: $|p\bullet| > 1$ implies $|\bullet t| = 1$
    - For all $p \in P$ and $t \in T$ with $(p, t) \in F$: $|\bullet t| > 1$ implies $|\bullet p| = 1$
    - There are no implicit places



- Note
  - Sufficiently expressive to model most process-related control-flow structures, sequences, concurrency, exclusive choices, etc.

Soundness

- Behavioural correctness criterion for WF-nets
  - Processes terminate in proper final state
  - Final state is indeed characterised unambiguously
  - All activities can contribute to process execution

- Recall: WF-net has initial place i and final place o
  - Overload notation and refer to i and o also as the markings that put one token into i and o, respectively, and no token in any other place

Soundness Definition

- A workflow system (PN, i) with workflow net $PN = (P, T, F)$ is *sound* iff
  - For every state M reachable from state i there exists a firing sequence leading from M to o *(option to complete)*:

$$\forall M \left(i \xrightarrow{*} M \right) \Rightarrow \left(M \xrightarrow{*} o \right)$$

  - State o is the only state reachable from state i with at least one token in place o *(proper completion)*:

$$\forall M \left(i \xrightarrow{*} M \wedge M \geq o \right) \Rightarrow M = o$$

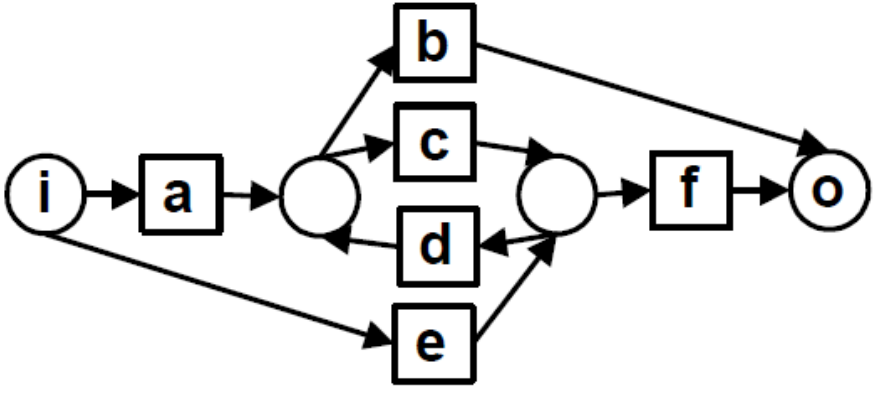
  - There are no dead transition in the workflow net in state i *(no dead transition):*

$$(\forall t \in T)\ \exists\, M, M' : i \xrightarrow{*} M \xrightarrow{t} M'$$

What can be rediscovered?

- Consider Loops

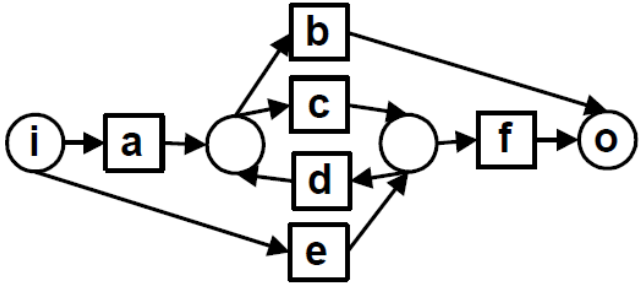
Original net:



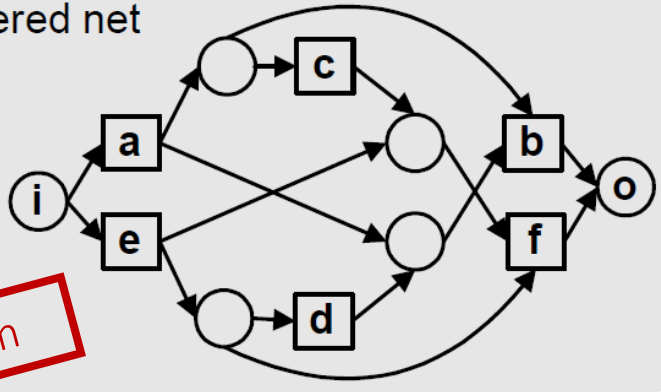
What can be rediscovered?

- Short loops are a problem

Original net:



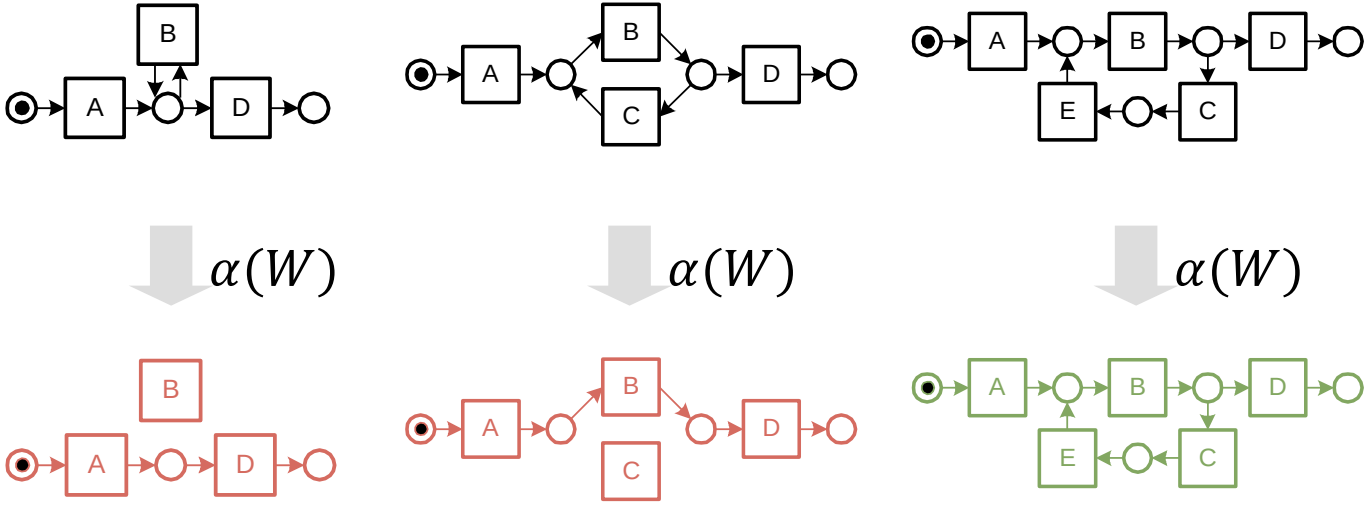
Discovered net



α Algorithm

Short Loops

- Short loops (length 1 or 2) pose an issue for rediscoverability by the α Algorithm:



What can be rediscovered?

The α Algorithm can rediscover

***sound structured workflow net without short loops**,*

if the event log is complete according to introduced notion

(based on direct success relationship)