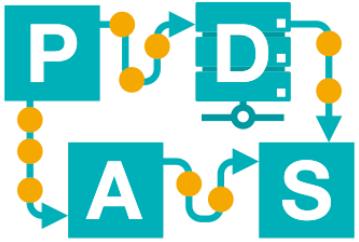
Conformance Checking

Footprint &Token-Based Replay

BPI-I 8

www.pads.rwth-aachen.de

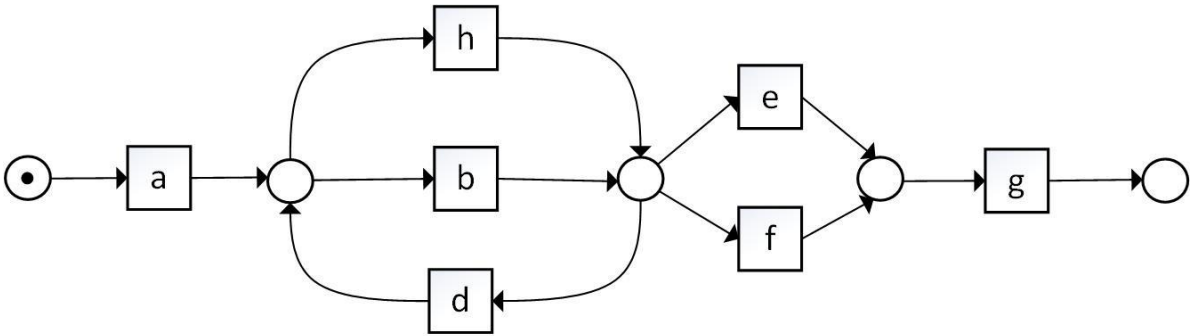


Chair of Process and Data Science

RWTH AACHEN UNIVERSITY

Exercise 1

- **Compute the footprint-based conformance for the event log L and the presented model.**
  - $L = [\langle a, b, f, e, g \rangle^{45}, \langle a, b, d, b, e, f, g \rangle^{5}, \langle a, h, e, f, g \rangle^{25}, \langle a, h, f, e, g \rangle^{25}]$

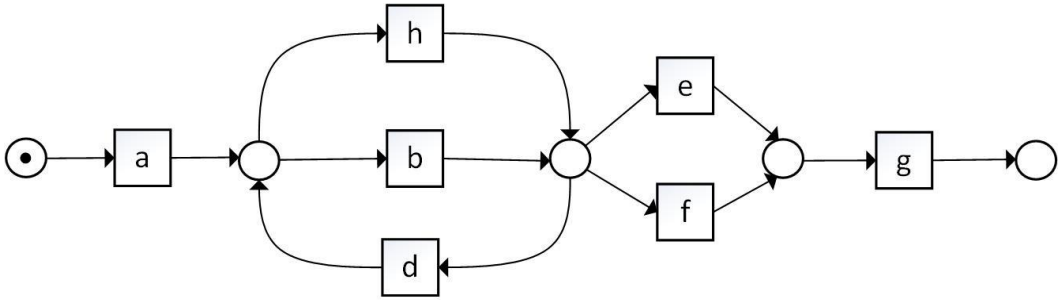


Exercise 1 Solution

- $L = [\langle a, b, f, e, g \rangle^{45}, \langle a, b, d, b, e, f, g \rangle^{5}, \langle a, h, e, f, g \rangle^{25}, \langle a, h, f, e, g \rangle^{25}]$

	a	b	d	e	f	g	h
a	#	→	#	#	#	#	→
b	←	#	||	→	→	#	#
d	#	||	#	#	#	#	#
e	#	←	#	#	||	→	←
f	#	←	#	||	#	→	←
g	#	#	#	←	←	#	#
h	←	#	#	→	→	#	#

Exercise 1 Solution



	a	b	d	e	f	g	h
a	#	→	#	#	#	#	→
b	←	#	||	→	→	#	#
d	#	||	#	#	#	#	||
e	#	←	#	#	#	→	←
f	#	←	#	#	#	→	←
g	#	#	#	←	←	#	#
h	←	#	||	→	→	#	#

Exercise 1 Solution

Event log

	a	b	d	e	f	g	h
a	#	→	#	#	#	#	→
b	←	#	||	→	→	#	#
d	#	||	#	#	#	#	#
e	#	←	#	#	||	→	←
f	#	←	#	||	#	→	←
g	#	#	#	←	←	#	#
h	←	#	#	→	→	#	#

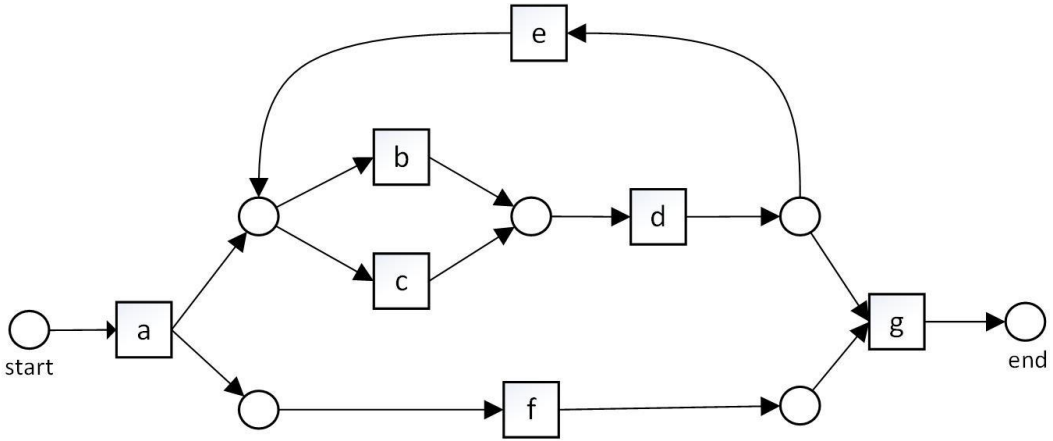
Process Model

	a	b	d	e	f	g	h
a	#	→	#	#	#	#	→
b	←	#	||	→	→	#	#
d	#	||	#	#	#	#	||
e	#	←	#	#	#	→	←
f	#	←	#	#	#	→	←
g	#	#	#	←	←	#	#
h	←	#	||	→	→	#	#

Footprint-based conformance: $1 - \frac{4}{49} = 0.918$

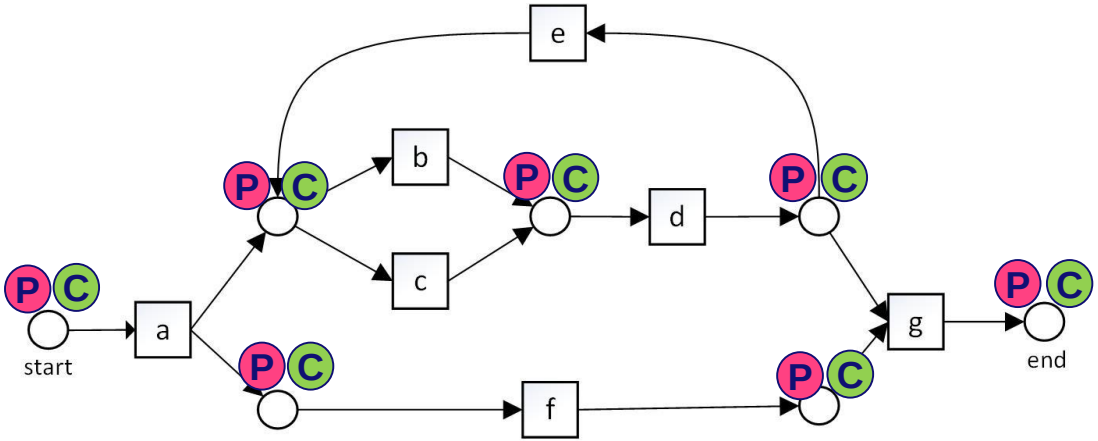
Exercise 2

- **Compute the token-based replay fitness for the event log L and the presented model.**
  - $L = [\langle a, b, d, f, g \rangle^{8}, \langle a, c, d, e, b, f, g \rangle^{4}, \langle a, b, c, d, f, g \rangle^{5}, \langle a, b, f, g \rangle^{1}, \langle b, d, f, g \rangle^{2}]$



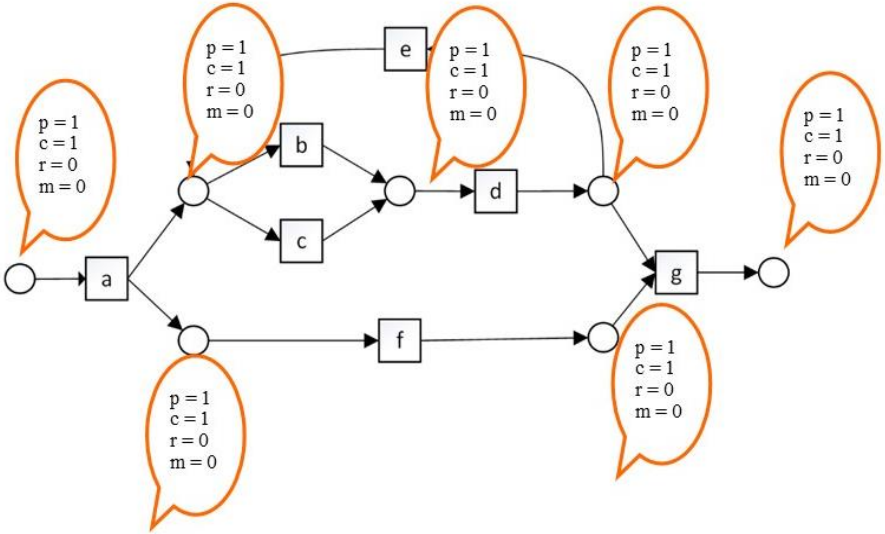
Exercise 2 Solution

$\langle a, b, d, f, g \rangle$



Exercise 2 Solution

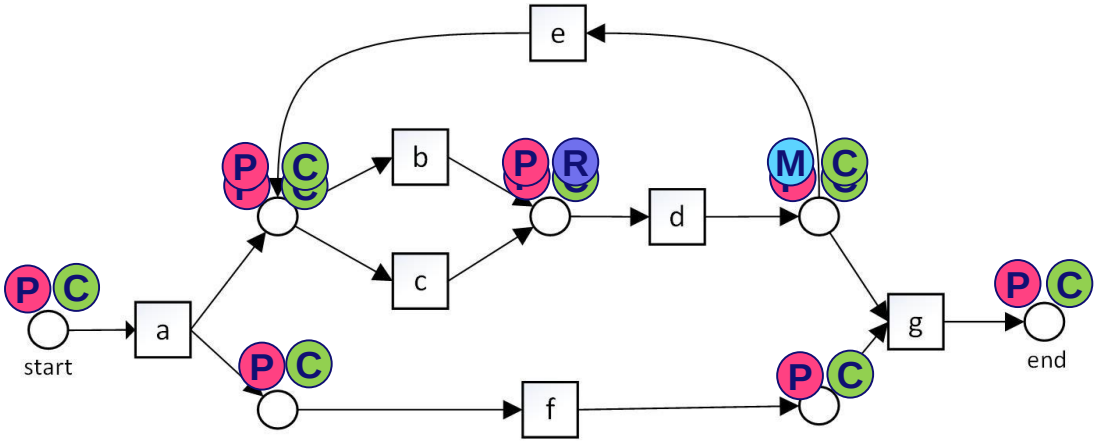
$\langle a, b, d, f, g \rangle$



$p = 7; c = 7; r = 0; m = 0$

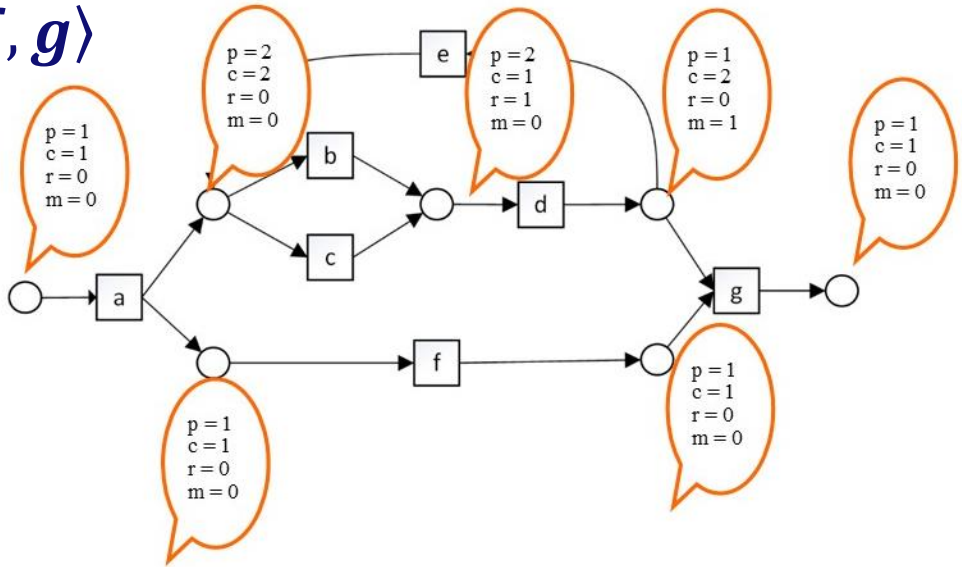
Exercise 2 Solution

$\langle a, c, d, e, b, f, g \rangle$



Exercise 2 Solution

$\langle a, c, d, e, b, f, g \rangle$

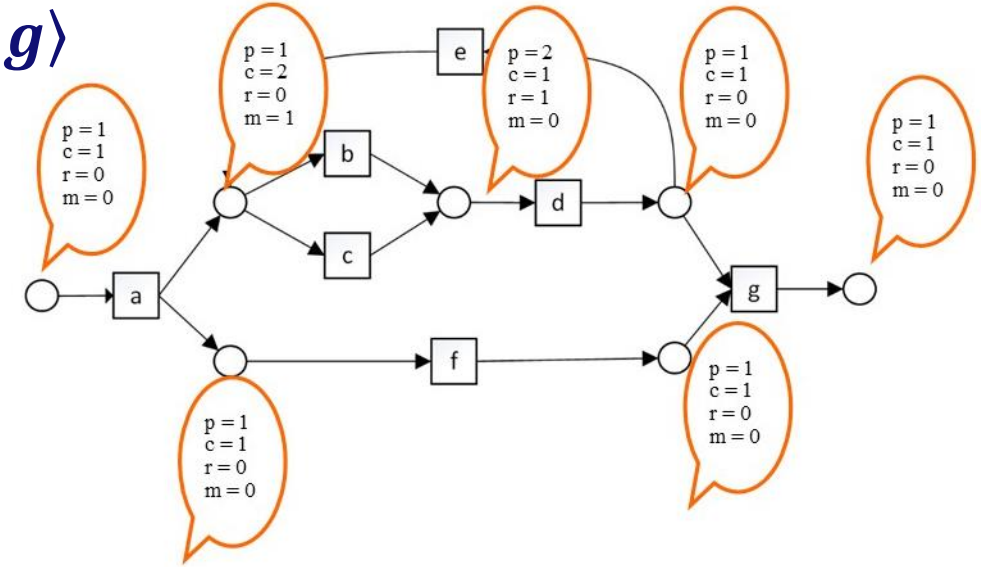


$p = 9; c = 9; r = 1; m = 1$

Chair of Process and Data Science

Exercise 2 Solution

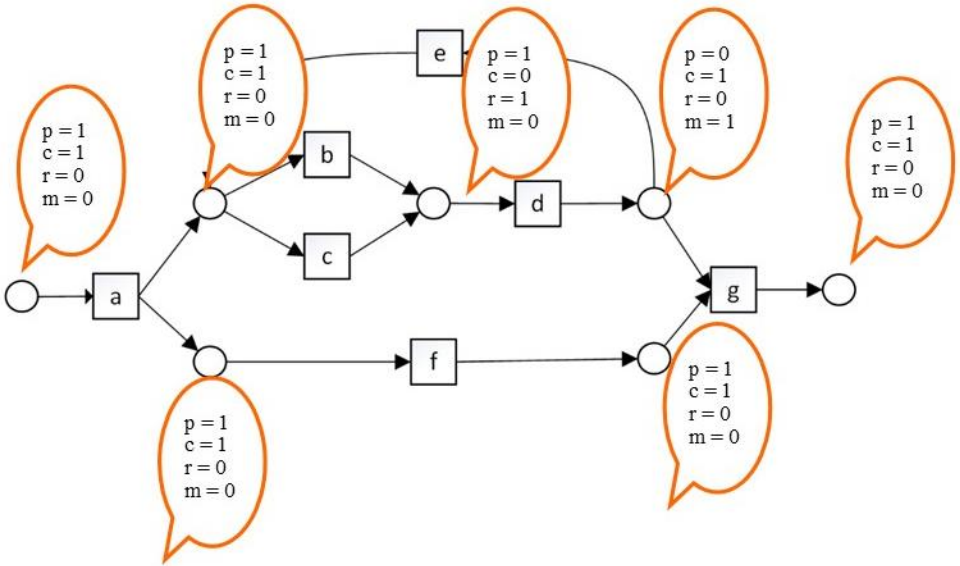
$\langle a, b, c, d, f, g \rangle$



$p = 8; c = 8; r = 1; m = 1$

Exercise 2 Solution

$\langle a, b, f, g \rangle$

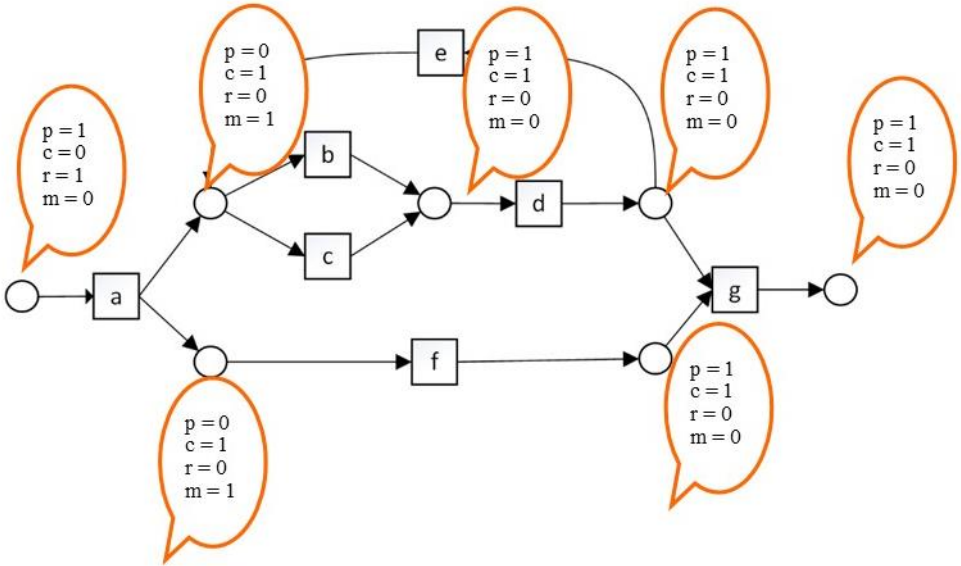


$p = 6; c = 6; r = 1; m = 1$

Chair of Process and Data Science

Exercise 2 Solution

$\langle b, d, f, g \rangle$



$p = 5; c = 6; r = 1; m = 2$

Exercise 2 Solution

- **Token-Based Fitness:**

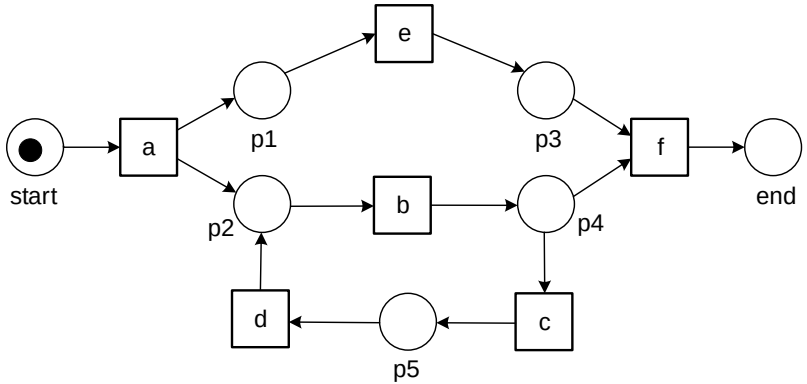
$$\mathbf{fitness}(L,N) = \frac{1}{2}\left(1 - \frac{\sum_{\sigma \in L} L(\sigma) \cdot m_{N,\sigma}}{\sum_{\sigma \in L} L(\sigma) \cdot c_{N,\sigma}}\right) + \frac{1}{2}\left(1 - \frac{\sum_{\sigma \in L} L(\sigma) \cdot r_{N,\sigma}}{\sum_{\sigma \in L} L(\sigma) \cdot p_{N,\sigma}}\right)$$

- $\sum_{\sigma \in L} L(\sigma) \cdot p_{N,\sigma} = (8 \cdot 7) + (1 \cdot 6) + (5 \cdot 8) + (2 \cdot 5) + (4 \cdot 9) = 148$
- $\sum_{\sigma \in L} L(\sigma) \cdot c_{N,\sigma} = (8 \cdot 7) + (1 \cdot 6) + (5 \cdot 8) + (2 \cdot 6) + (4 \cdot 9) = 150$
- $\sum_{\sigma \in L} L(\sigma) \cdot r_{N,\sigma} = (8 \cdot 0) + (1 \cdot 1) + (5 \cdot 1) + (2 \cdot 1) + (4 \cdot 1) = 12$
- $\sum_{\sigma \in L} L(\sigma) \cdot m_{N,\sigma} = (8 \cdot 0) + (1 \cdot 1) + (5 \cdot 1) + (2 \cdot 2) + (4 \cdot 1) = 14$

$$\frac{1}{2}\left(1 - \frac{14}{150}\right) + \frac{1}{2}\left(1 - \frac{12}{148}\right) = 0.912$$

Exercise 3

- **Compute the footprint and token-based replay fitness for the event log L and the presented model.**
  - $L = [\langle a, b, e, f \rangle^2, \langle a, c, d, f \rangle^3]$



Exercise 3 Solution

Footprint

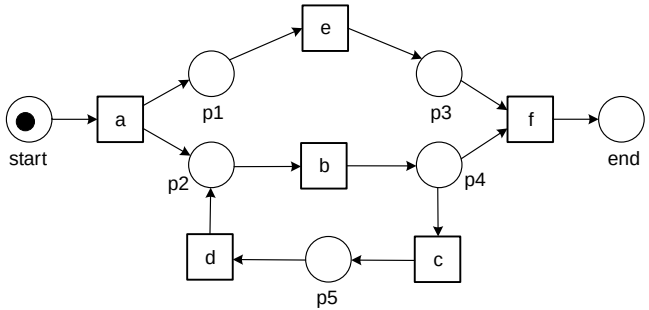
From the log

L	a	b	c	d	e	f
a	#	→	→	#	#	#
b	←	#	#	#	→	#
c	←	#	#	→	#	#
d	#	#	←	#	#	→
e	#	←	#	#	#	→
f	#	#	#	←	←	#

From the model

M	a	b	c	d	e	f
a	#	→	#	#	→	#
b	←	#	→	←	||	→
c	#	←	#	→	||	#
d	#	→	←	#	||	#
e	←	||	||	||	#	→
f	#	←	#	#	←	#

$$L = [\langle a, b, e, f \rangle^2, \langle a, c, d, f \rangle^3]$$



$$1 - \frac{18}{36} = 0.5$$

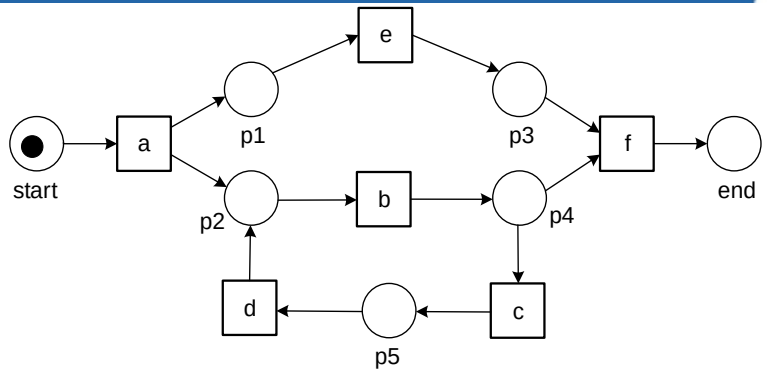
Chair of Process and Data Science

Exercise 3 Solution

Token-based Replay

- $L = [\langle a, b, e, f \rangle^2, \langle a, c, d, f \rangle^3]$

p = ?
c = ?
m = ?
r = ?



First trace?

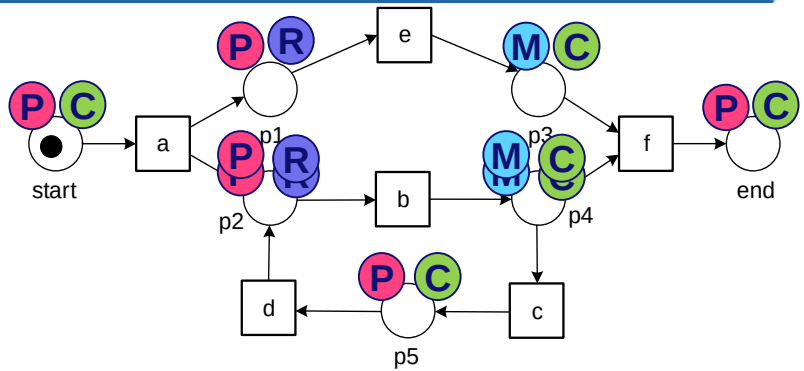
Trace	Frequency	P	R	C	M	P <all>	R <all>	C <all>	M <all>
<a, b, e, f>	2	6	0	6	0	12	0	12	0
<a, c, d, f>	3								
total									

Why?

Exercise 3 Solution

Token-based Replay

L = [<a, b, e, f>^{2}, <a, c, d, f>^{3}]
Second trace?

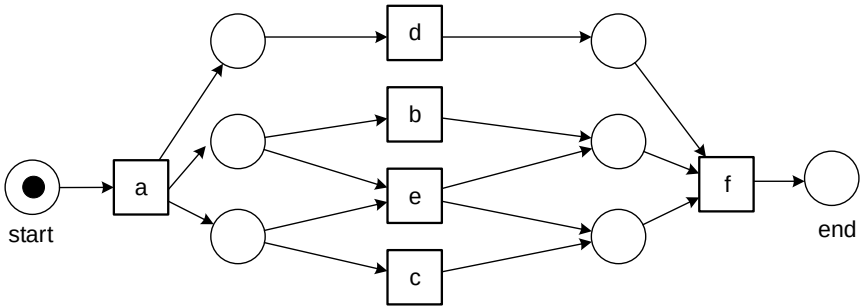


$$\text{fitness} = \frac{1}{2}\left(1 - \frac{9}{30}\right) + \frac{1}{2}\left(1 - \frac{9}{30}\right) = 0.7$$

Trace	Frequency	P	R	C	M	P <all>	R <all>	C <all>	M <all>
<a, b, e, f>	2	6	0	6	0	12	0	12	0
<a, c, d, f>	3	6	3	6	3	18	9	18	9
total						**30**	**9**	**30**	**9**

Exercise 4

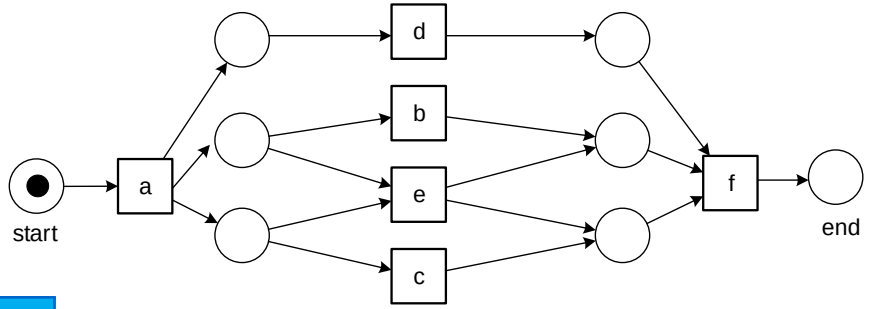
- **Compute the footprint and token-based replay fitness for the event log L and the presented model.**
  - $L = [\langle a, c, d, f \rangle, \langle a, c, b, d, f \rangle^2, \langle a, b, c, f \rangle, \langle a, e, f \rangle^2]$



Exercise 4 Solution

Footprint

- $L = [\langle a, c, d, f \rangle, \langle a, c, b, d, f \rangle^2, \langle a, b, c, f \rangle, \langle a, e, f \rangle^2]$



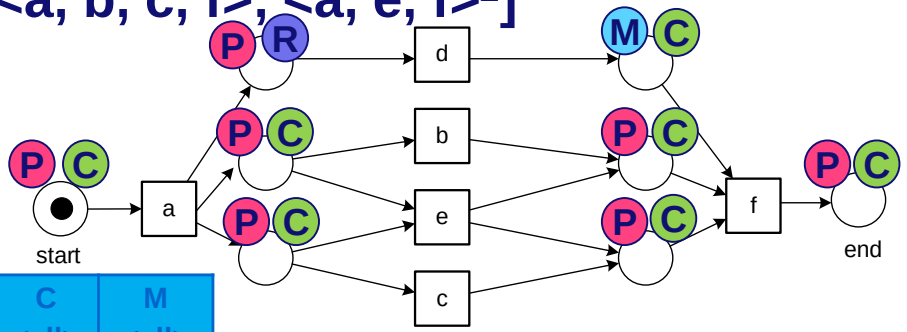
M : L	a	b	c	d	e	f
a	#	→	→	→ : #	→	#
b	←	#	||	|| : →	#	→ : #
c	←	||	#	|| : →	#	→
d	← : #	|| : ←	|| : ←	#	|| : #	→
e	←	#	#	|| : #	#	→
f	#	← : #	←	←	←	#

$$1 - \frac{10}{36} = 0.72$$

Exercise 4 Solution

Token-based Replay

- L =[<a, c, d, f>, <a, c, b, d, f>^{2}, <a, b, c, f>, <a, e, f>^{2}]

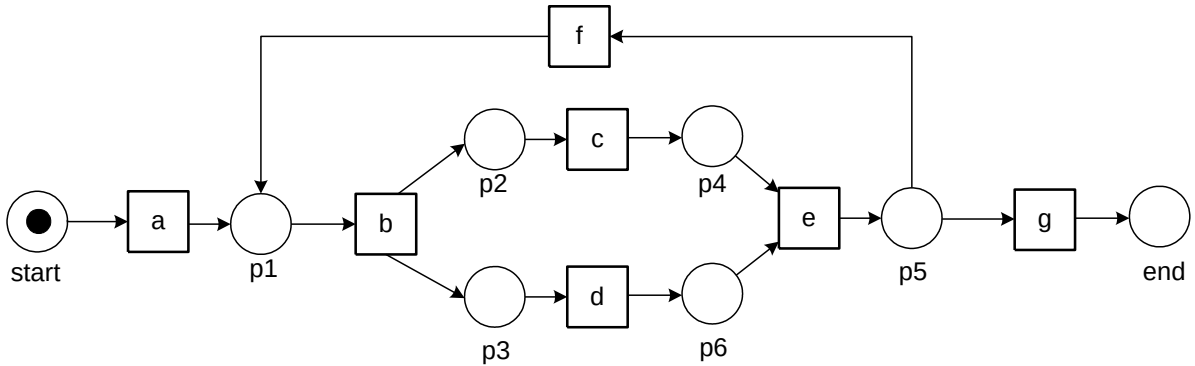


Trace	Frequency	P	R	C	M	P <all>	R <all>	C <all>	M <all>
<a, c, d, f>	1	7	1	7	1	7	1	7	1
<a, c, b, d, f>	2	8	0	8	0	16	0	16	0
<a, b, c, f>	1	7	1	7	1	7	1	7	1
<a, e, f>	2	7	1	7	1	14	2	14	2
total						44	4	44	4

$$\text{fitness} = \frac{1}{2}\left(1 - \frac{4}{44}\right) + \frac{1}{2}\left(1 - \frac{4}{44}\right) = 0.9$$

Exercise 5

- **Compute the footprint and token-based replay fitness for the event log L and the presented model.**
  - $L = [\langle a, b, c, d, e, f, b, d, c, g \rangle, \langle a, b, d, c, e, g \rangle, \langle a, b, c, d, e, f, c, d, e, f, b, d, e, g \rangle]$



Exercise 5 Solution

Footprint

L	a	b	c	d	e	f	g
a	#	→	#	#	#	#	#
b	←	#	→	→	#	←	#
c	#	←	#	||	→	←	→
d	#	←	||	#	→	#	#
e	#	#	←	←	#	→	→
f	#	→	→	#	←	#	#
G	#	#	←	#	←	#	#

$$1 - {}^{4}/_{49} = 0.9184$$

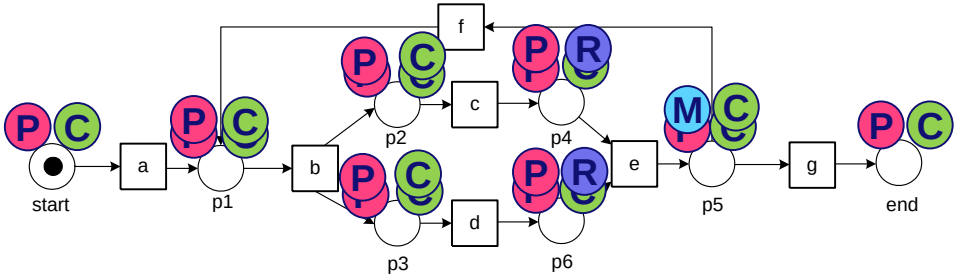
M	a	b	c	d	E	f	g
a	#	→	#	#	#	#	#
b	←	#	→	→	#	←	#
c	#	←	#	||	→	#	#
d	#	←	||	#	→	#	#
e	#	#	←	←	#	→	→
f	#	→	#	#	←	#	#
g	#	#	#	#	←	#	#

Chair of Process and Data Science

Exercise 5 Solution

Token-based Replay

- **L = [<a, b, c, d, e, f, b, d, c, g>, <a, b, d, c, e, g>,<a, b, c, d, e, f, c, d, e, f, b, d, e, g>]**

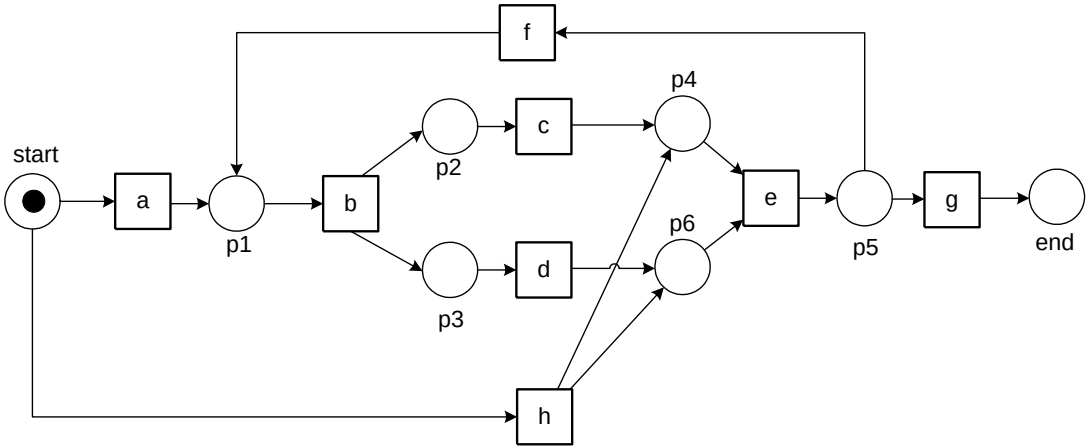


Trace	Frequency	P	R	C	M	P <all>	R <all>	C <all>	M <all>
<a, b, c, d, e, f, b, d, c, g>	1	13	2	12	1	13	2	12	1
<a, b, d, c, e, g>	1	8	0	8	0	8	0	8	0
<a, b, c, d, e, f, c, d, e, f, b, d, e, g>	1	17	2	18	3	17	2	18	3
total						**38**	**4**	**38**	**4**

$$\textbf{fitness} = \frac{1}{2}\left(1 - \frac{4}{38}\right) + \frac{1}{2}\left(1 - \frac{4}{38}\right) = 0.8947$$

Exercise 6

- **Compute the footprint and token-based replay fitness for the event log L and the presented model.**
  - $L = [\langle a, b, c, e, f, b, d, c, e, g \rangle, \langle a, b, e, g \rangle, \langle h, e, g \rangle]$

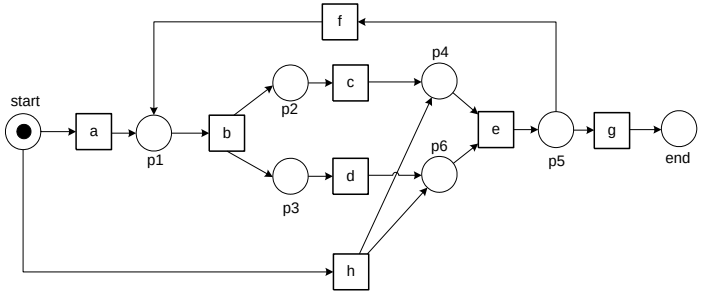


Exercise 6 Solution

Footprint

$$L = [\langle a, b, c, e, f, b, d, c, e, g \rangle, \langle a, b, e, g \rangle, \langle h, e, g \rangle]$$

M : L	a	b	c	d	e	f	g	h
a	#	→	#	#	#	#	#	#
b	←	#	→	→	# : →	←	#	#
c	#	←	#	|| : ←	→	#	#	#
d	#	←	|| : →	#	→ : #	#	#	#
e	#	# : ←	←	← : #	#	→	→	←
f	#	→	#	#	←	#	#	#
g	#	#	#	#	←	#	#	#
h	#	#	#	#	→	#	#	#



Footprint-based Conformance:

$$1 - \frac{6}{64} = 0.906$$

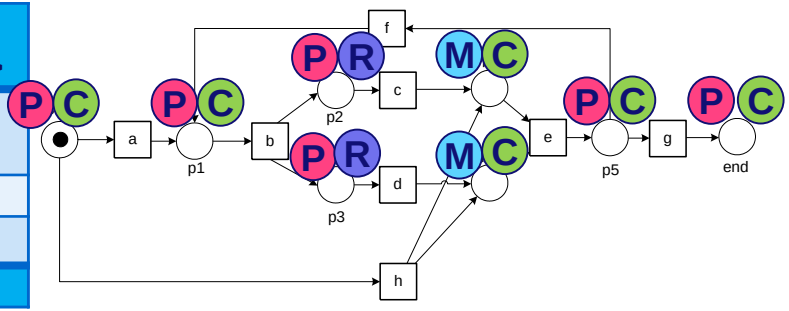
Chair of Process and Data Science

Exercise 6 Solution

Token-based Replay

L = [<a, b, c, e, f, b, d, c, e, g>,<a, b, e, g>, <h, e, g>]

Trace	Frequency	P	R	C	M	P <all>	R <all>	C <all>	M <all>
<a, b, c, e, f, b, d, c, e, g>	1	13	1	13	1	13	1	13	1
<a, b, e, g>	1	6	2	6	2	6	2	6	2
<h, e, g>	1	5	0	5	0	5	0	5	0
total						24	3	24	3



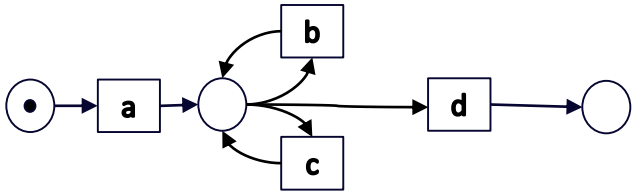
$$\text{fitness} = \frac{1}{2}\left(1 - \frac{3}{24}\right) + \frac{1}{2}\left(1 - \frac{3}{24}\right) = 0.875$$

Exercise 7

- **Design a sound workflow net using all the activities $\{a, b, c, d\}$ and only those activities such that it can replay the following trace but the fitness score using footprint table is below 0.7.**
  - $\langle a, b, c, d \rangle$
  - **What will be the token-based replay fitness, higher or less than footprint? Why?**

Exercise 7 Solution

$\langle a, b, c, d \rangle$



M : L	a	b	c	d
a	#	→	→:#	→:#
b	←	||:#	||:→	→:#
c	←:#	||:←	||:#	→
d	←:#	←:#	←	#

$p = 5$
$c = 5$
$m = 0$
$r = 0$

Footprint-Based Conformance: $1 - \frac{10}{16} = 0.375$

Token-Based Replay Fitness: $\frac{1}{2}\left(1 - \frac{0}{5}\right) + \frac{1}{2}\left(1 - \frac{0}{5}\right) = 1.00$