

CS 4072 - Topics in CS Process Mining

Lecture # 14

April 11, 2022

Spring 2022

FAST - NUCES, CFD Campus

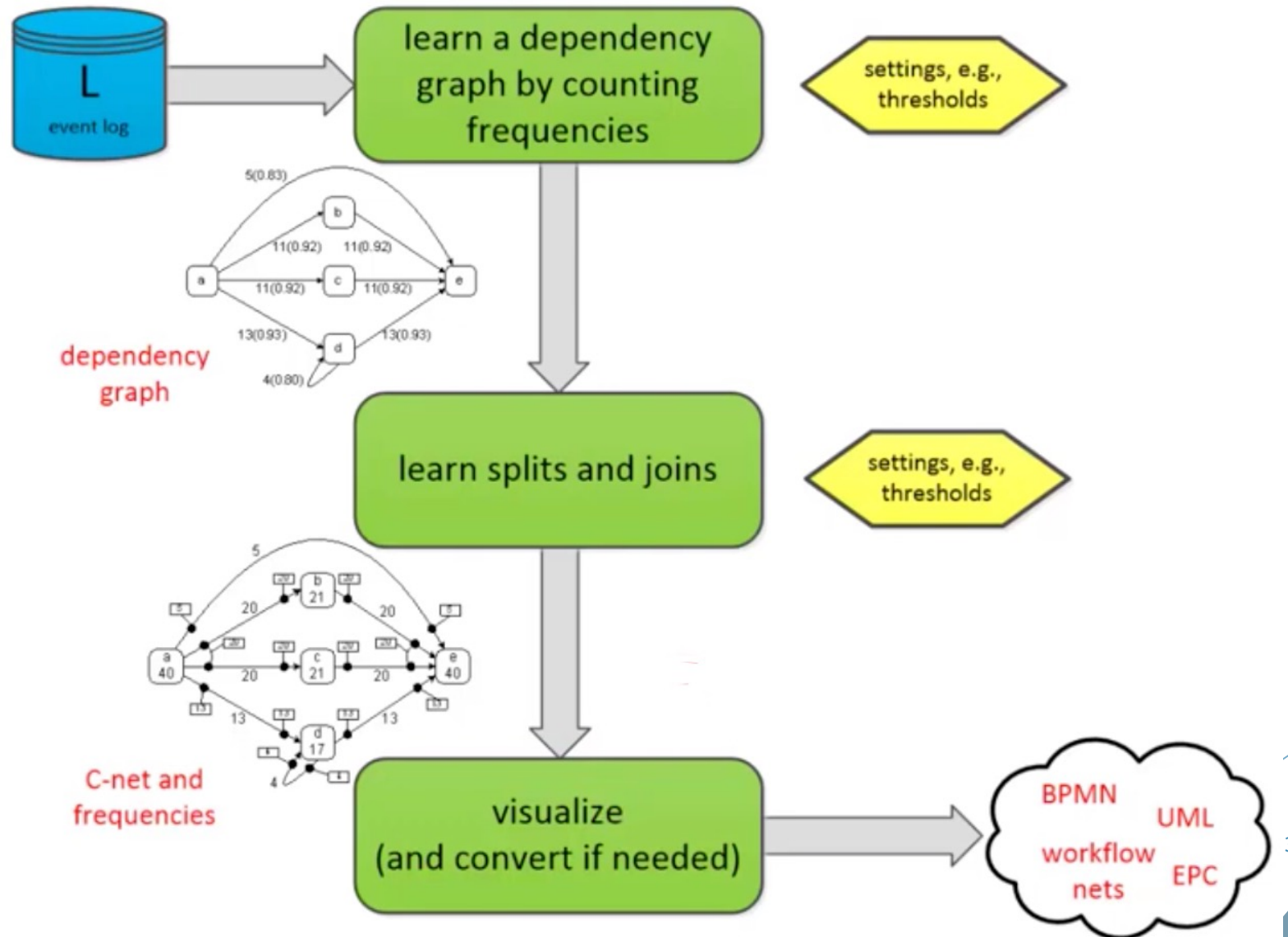
Dr. Rabia Maqsood

rabia.maqsood@nu.edu.pk

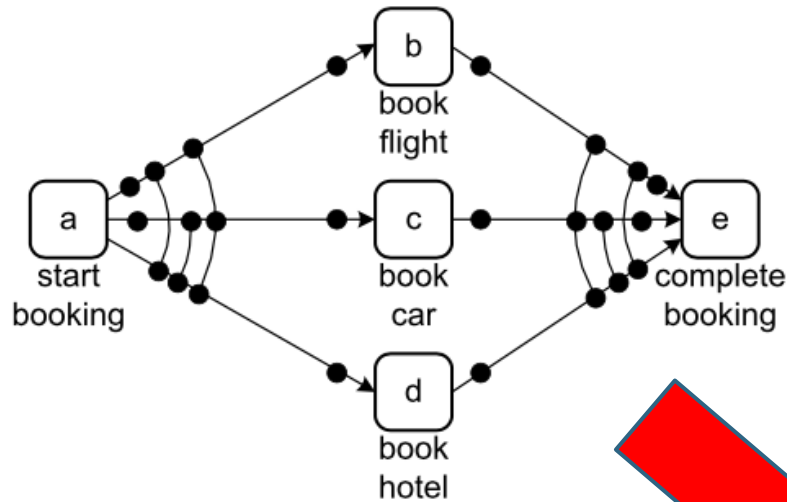
Today's Topics

- ▶ Heuristic Mining
 - ▶ Learning Dependency Graph
 - ▶ Learning Causal Nets and perform annotations

Heuristic Mining: two main phases

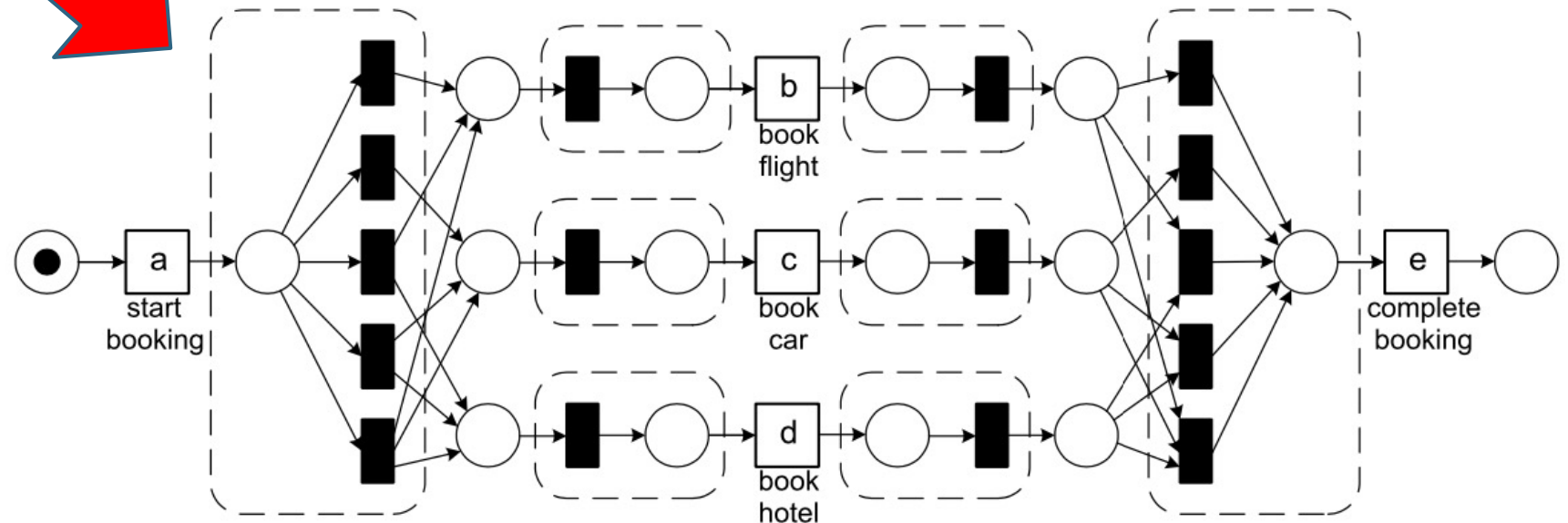
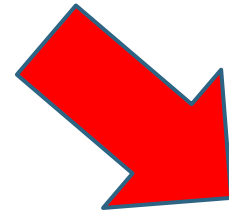


Relating C-nets to WF-nets

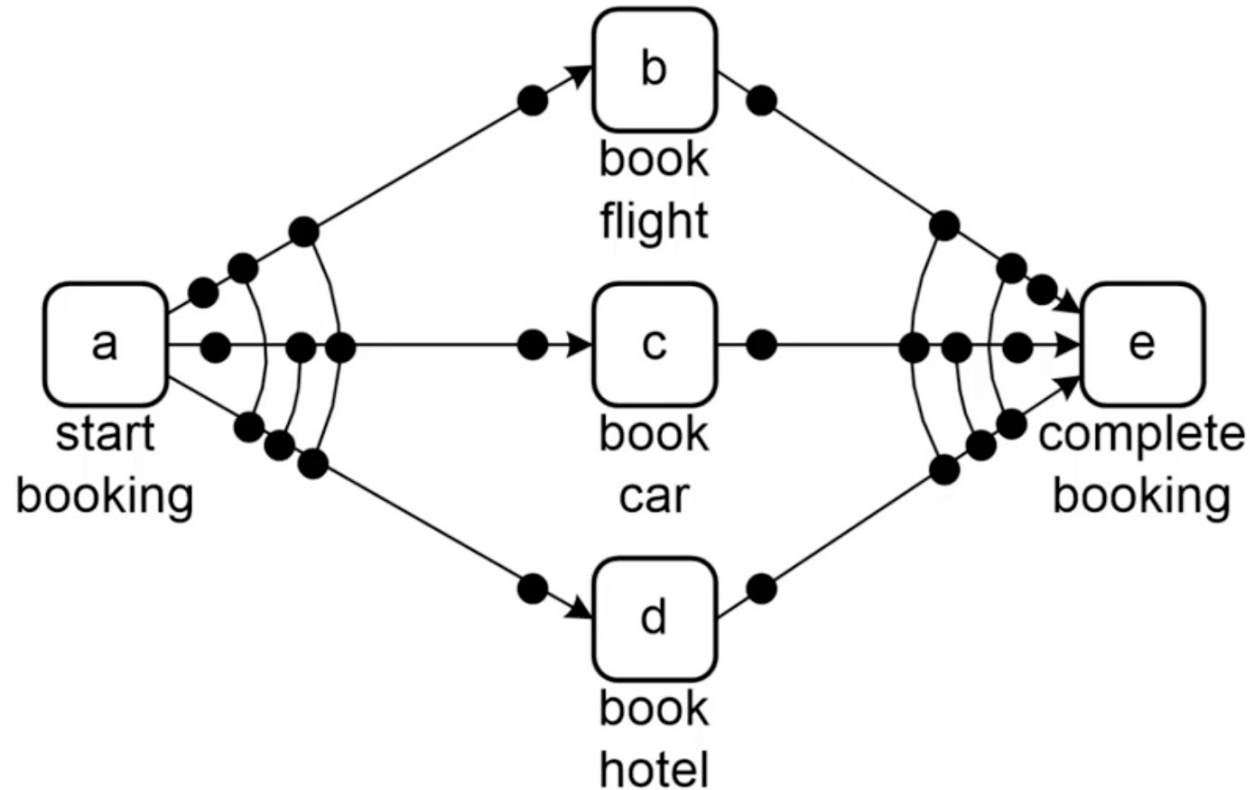


WF-net may have deadlocks, leave tokens behind, etc.

Only valid binding sequences are considered in C-nets.



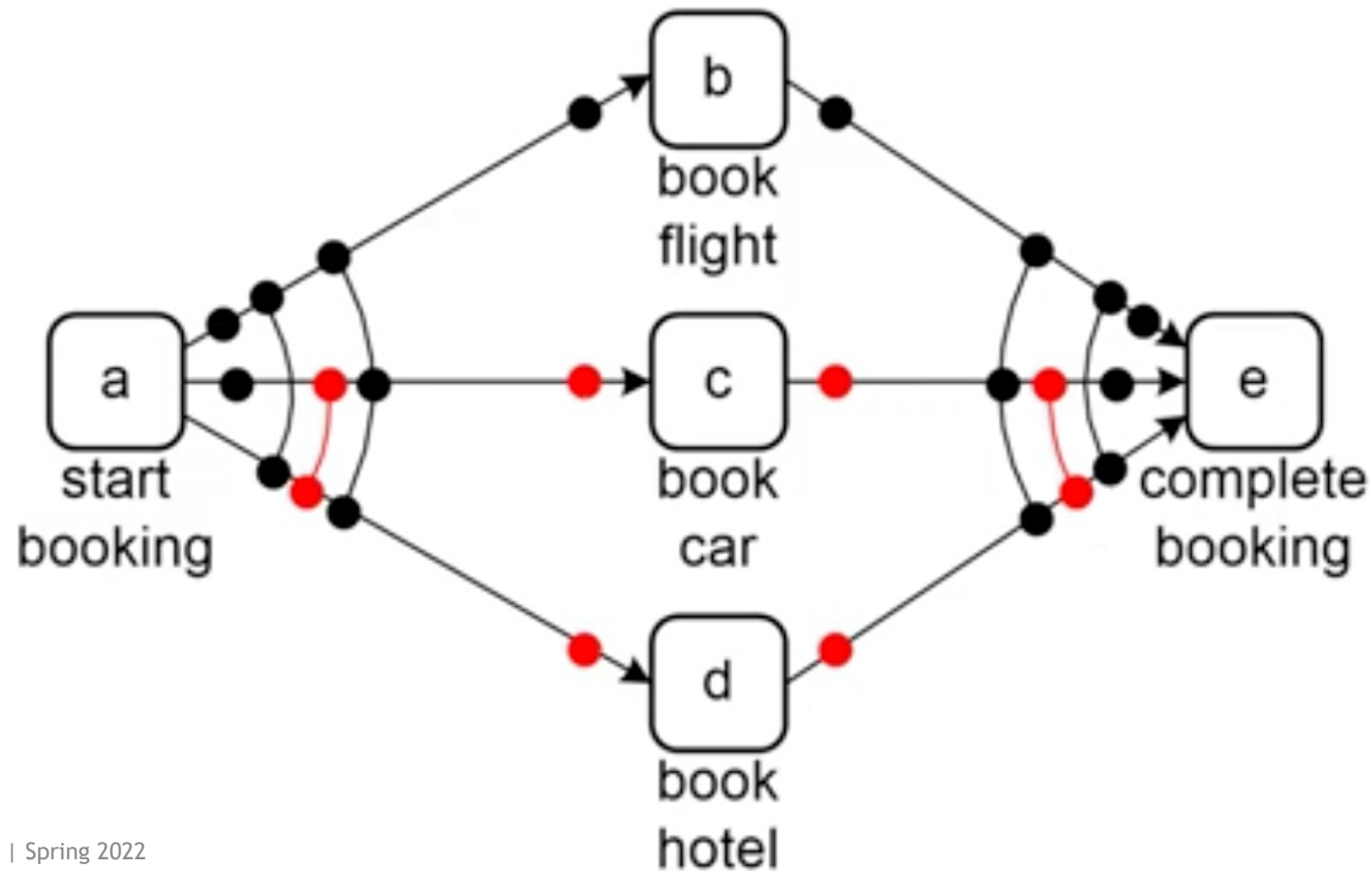
How many valid binding sequences?



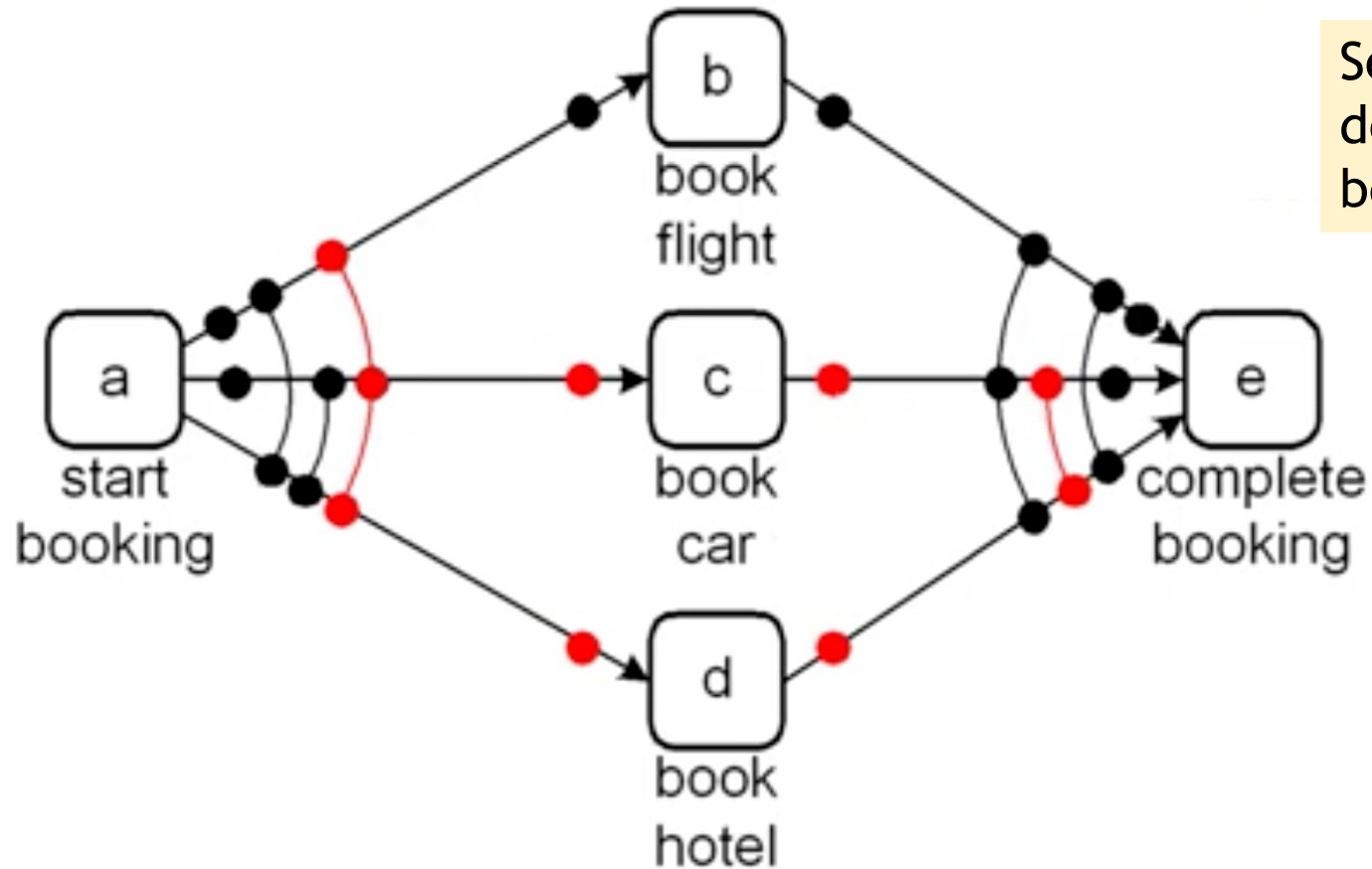
12 valid sequences

abe
ace
abde
adbe
acde
adce
abcde
abdce
acbde
acdbe
adbce
adcbe

Example **valid** binding sequence
generating trace **adce**

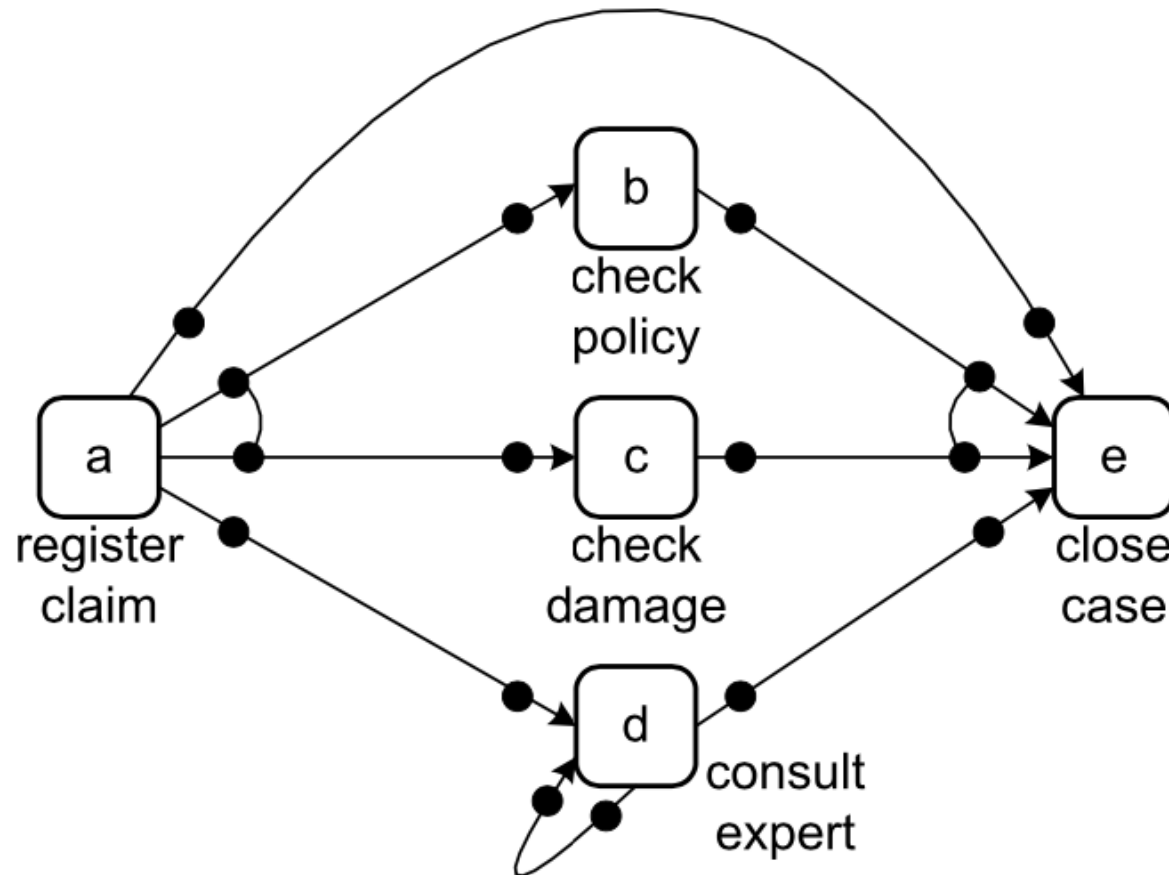


Example **invalid** binding sequence



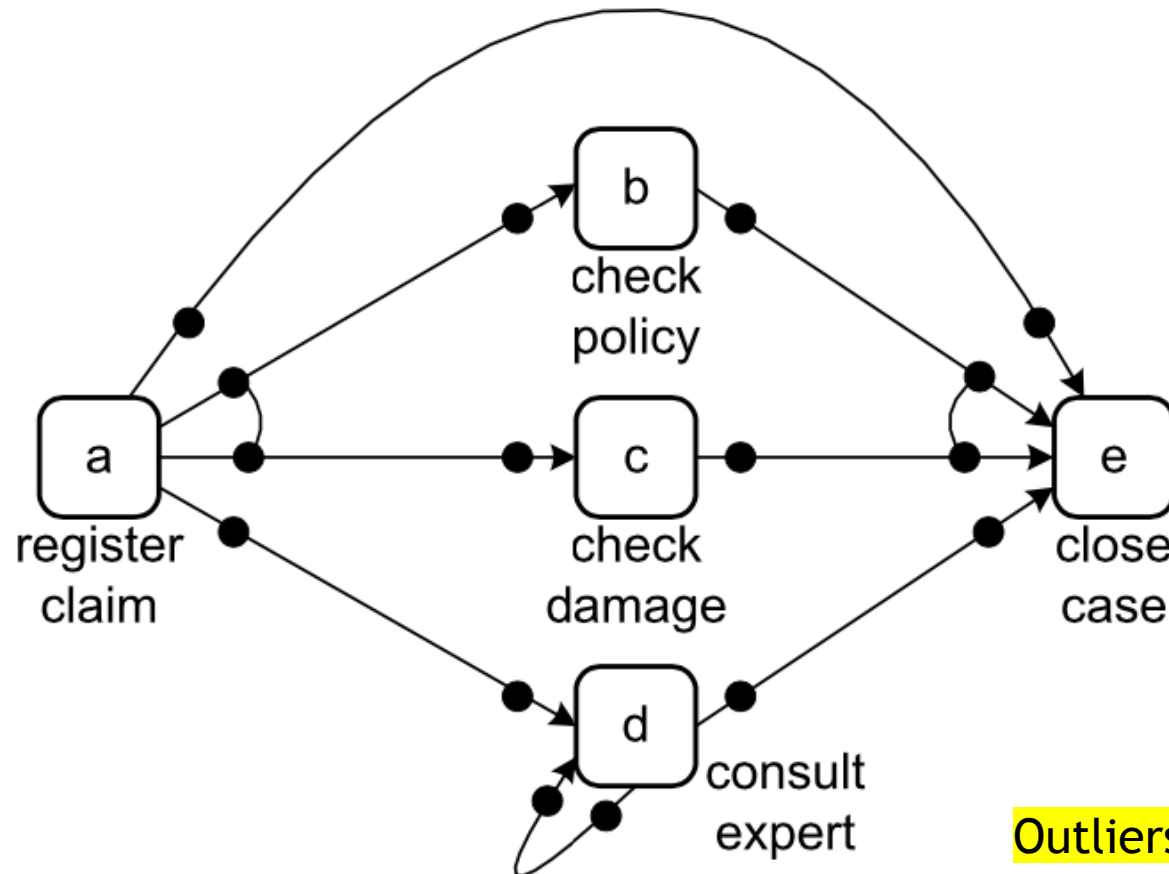
Semantics are declarative:
deadlocking or livelocking
behaviors are ignored!

Running example of a C-net



$L = [\langle a, e \rangle^5, \langle a, b, c, e \rangle^{10}, \langle a, c, b, e \rangle^{10}, \langle a, b, e \rangle^1, \langle a, c, e \rangle^1, \langle a, d, e \rangle^{10}, \langle a, d, d, e \rangle^2, \langle a, d, d, d, e \rangle^1]$

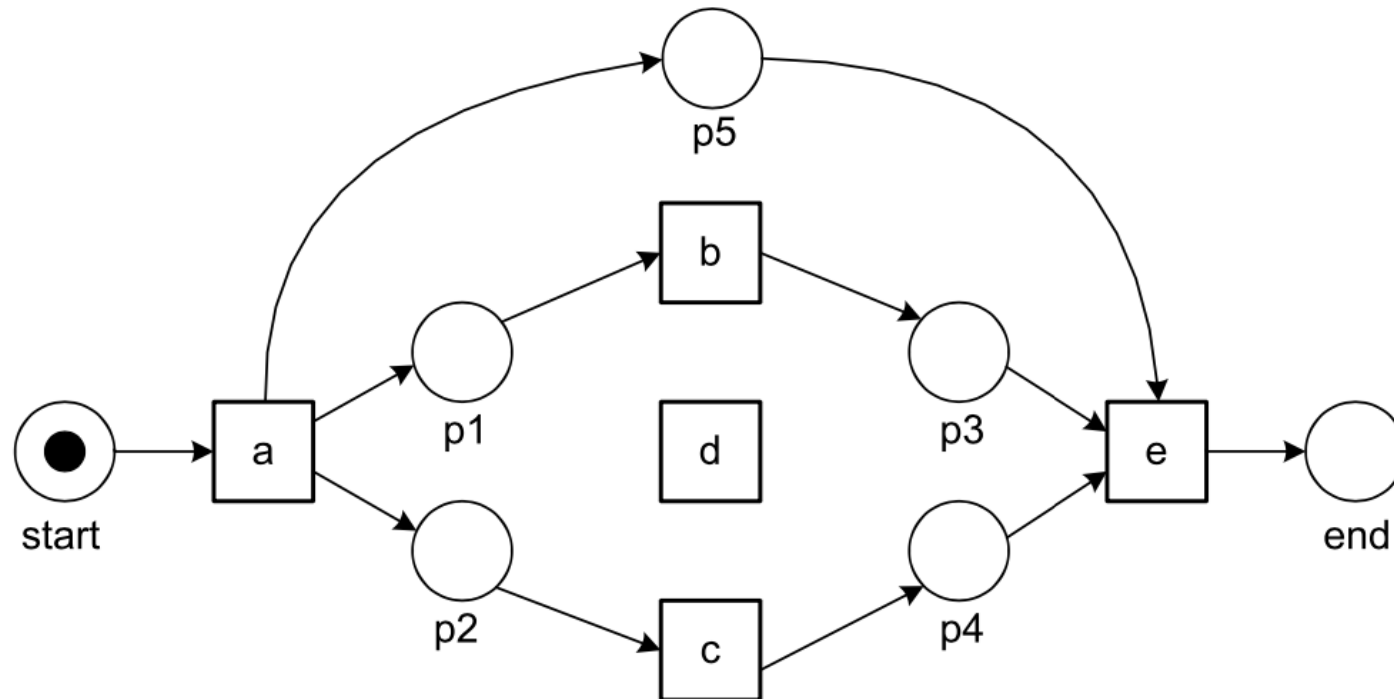
Running example of a C-net



Outliers not in the model

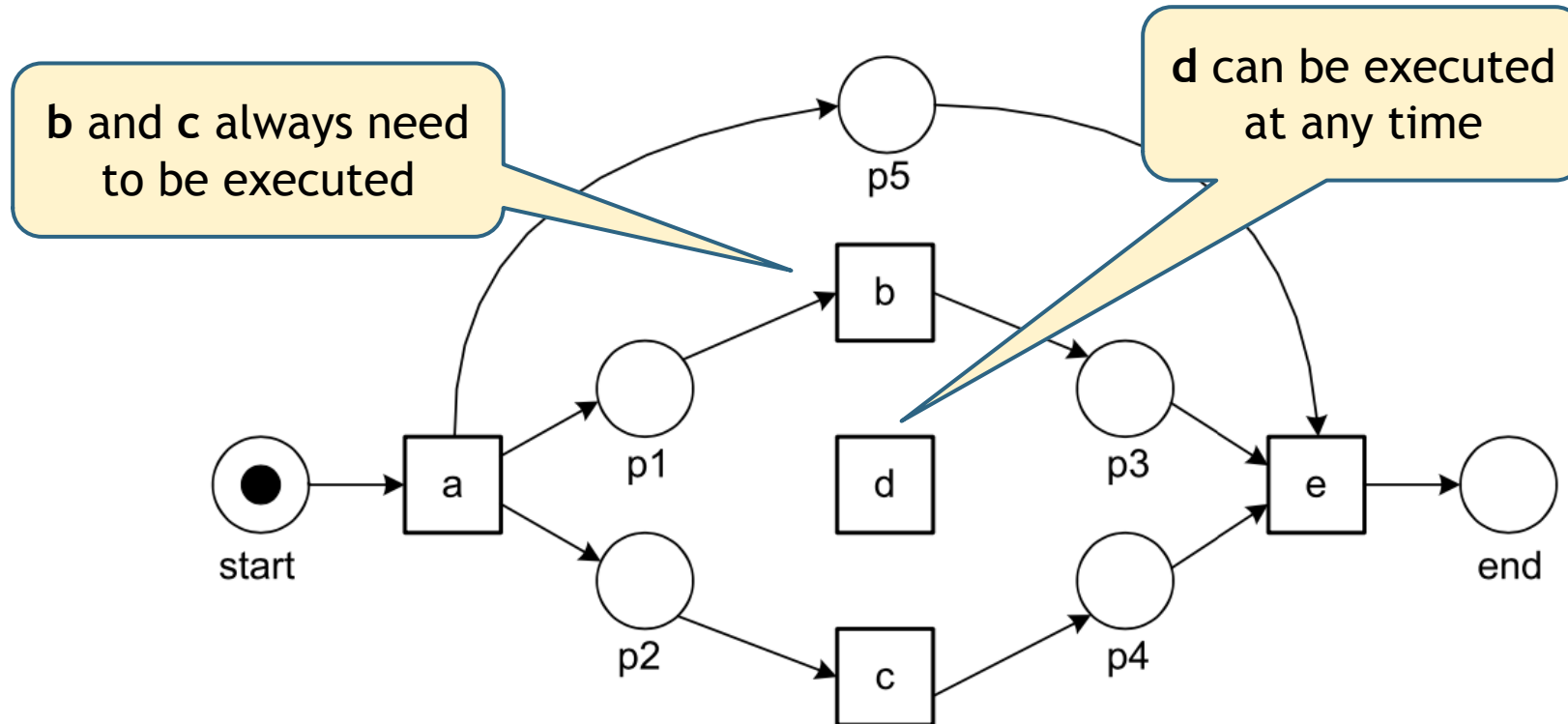
$L = [\langle a, e \rangle^5, \langle a, b, c, e \rangle^{10}, \langle a, c, b, e \rangle^{10}, \langle a, b, e \rangle^1, \langle a, c, e \rangle^1, \langle a, d, e \rangle^{10}, \langle a, d, d, e \rangle^2, \langle a, d, d, d, e \rangle^1]$

Limitations of the Alpha algorithm on such event logs

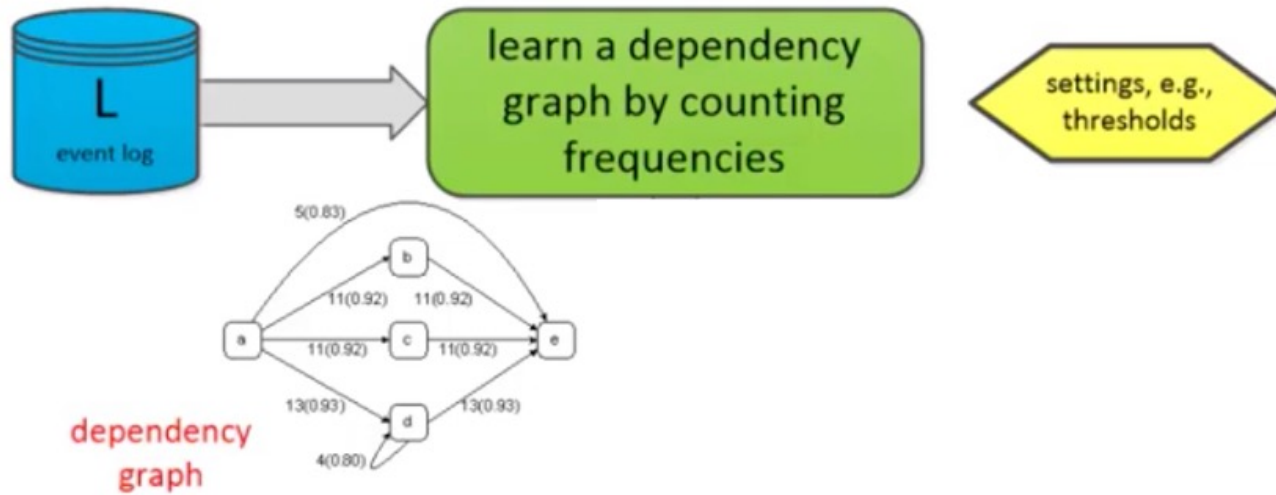


$L = [\langle a, e \rangle^5, \langle a, b, c, e \rangle^{10}, \langle a, c, b, e \rangle^{10}, \langle a, b, e \rangle^1, \langle a, c, e \rangle^1, \langle a, d, e \rangle^{10}, \langle a, d, d, e \rangle^2, \langle a, d, d, d, e \rangle^1]$

Limitations of the Alpha algorithm on such event logs



$L = [\langle a, e \rangle^5, \langle a, b, c, e \rangle^{10}, \langle a, c, b, e \rangle^{10}, \langle a, b, e \rangle^1, \langle a, c, e \rangle^1, \langle a, d, e \rangle^{10}, \langle a, d, d, e \rangle^2, \langle a, d, d, d, e \rangle^1]$

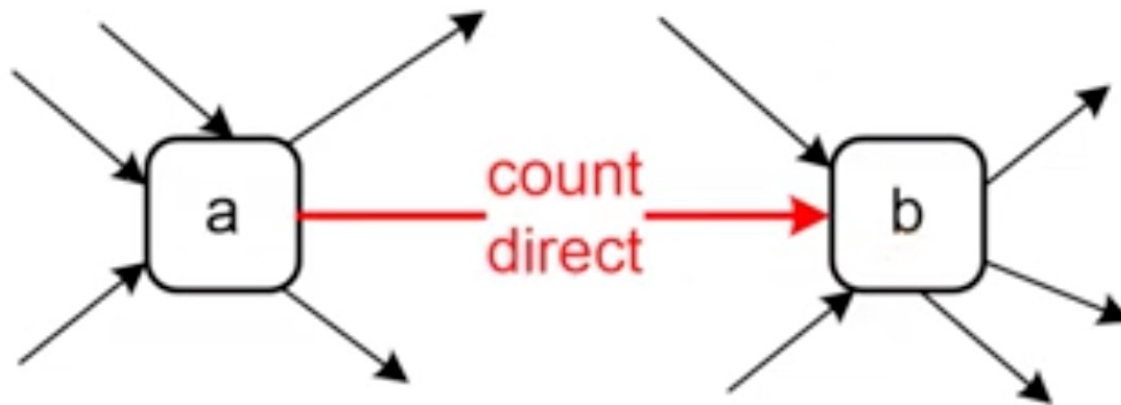


Learning Dependency Graph

First step in the Heuristic Mining

Frequencies matter!

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$



Counting Direct Succession

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$L = [\langle a, e \rangle^5, \langle a, b, c, e \rangle^{10}, \langle a, c, b, e \rangle^{10}, \langle a, b, e \rangle^1, \langle a, c, e \rangle^1, \langle a, d, e \rangle^{10}, \langle a, d, d, e \rangle^2, \langle a, d, d, d, e \rangle^1]$

$ >_L $	a	b	c	d	e
a	0	11	11	13	5
b	0	0	10	0	11
c	0	10	0	0	11
d	0	0	0	4	13
e	0	0	0	0	0

Counting Direct Succession

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$L = [\langle a, e \rangle^5, \langle a, b, c, e \rangle^{10}, \langle a, c, b, e \rangle^{10}, \langle a, b, e \rangle^1, \langle a, c, e \rangle^1, \langle a, d, e \rangle^{10}, \langle a, d, d, e \rangle^2, \langle a, d, d, d, e \rangle^1]$

$ >_L $	a	b	c	d	e
a	0	11	11	true	5
b	0	0	10	false	11
c	0	10	0	false	11
d	0	0	0	true	13
e	0	0	0	false	0

Information when frequencies are ignored

Dependency measure: taking into account concurrency

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

Direct Succession

Dependency Measure

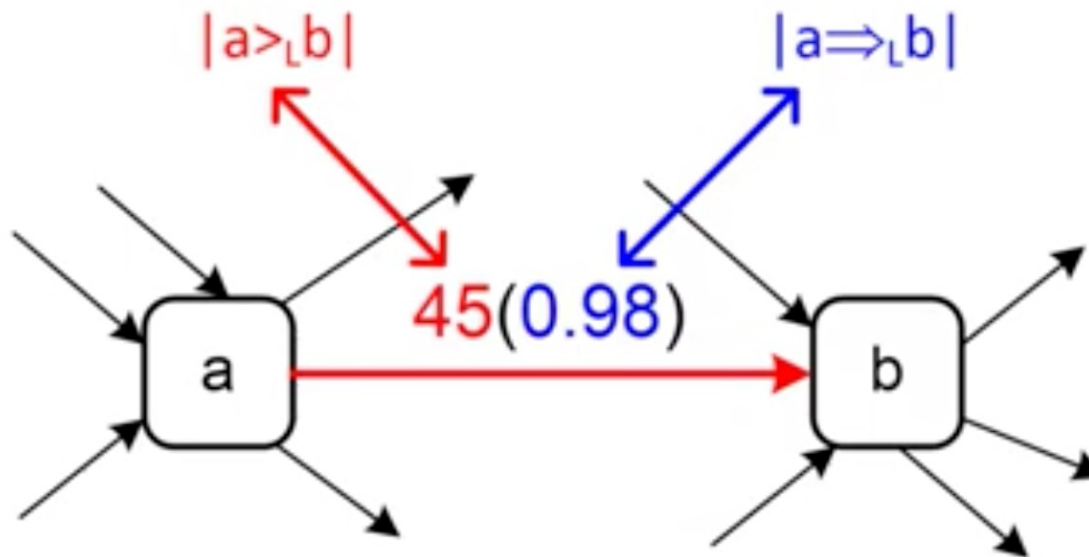
$|a \Rightarrow_L b|$ is the value of the dependency relation between a and b :

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$

Two values

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$



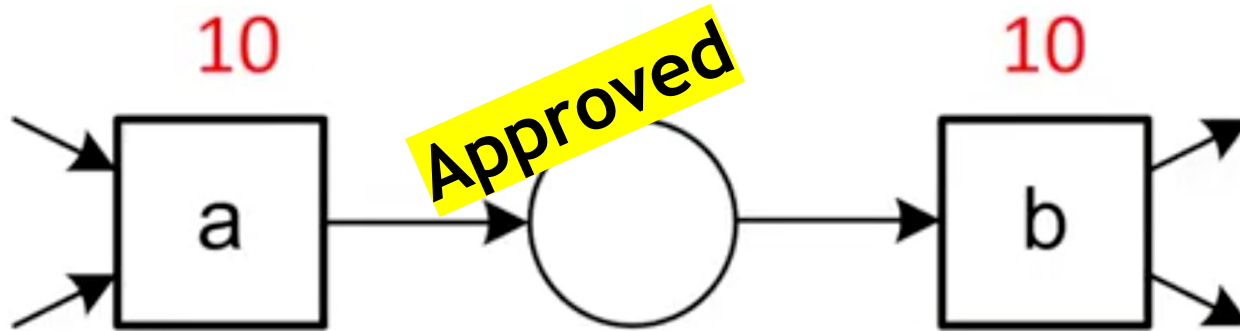
Both measures should be above some predefined thresholds!

Otherwise, no causality!

Sequence Pattern

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$



Approved

$$|a >_L b| = 10$$

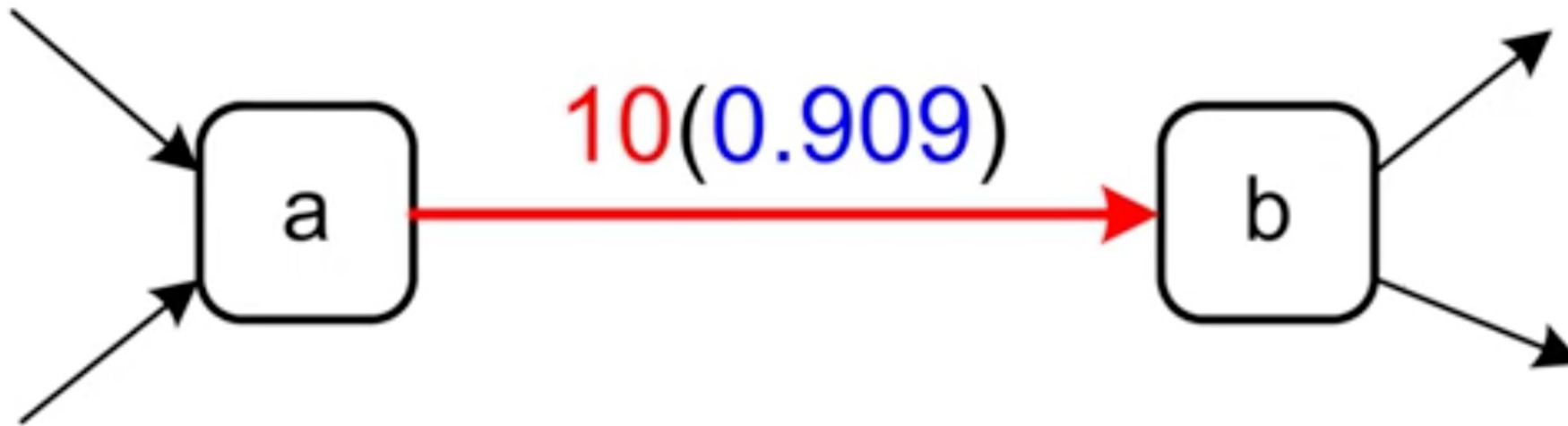
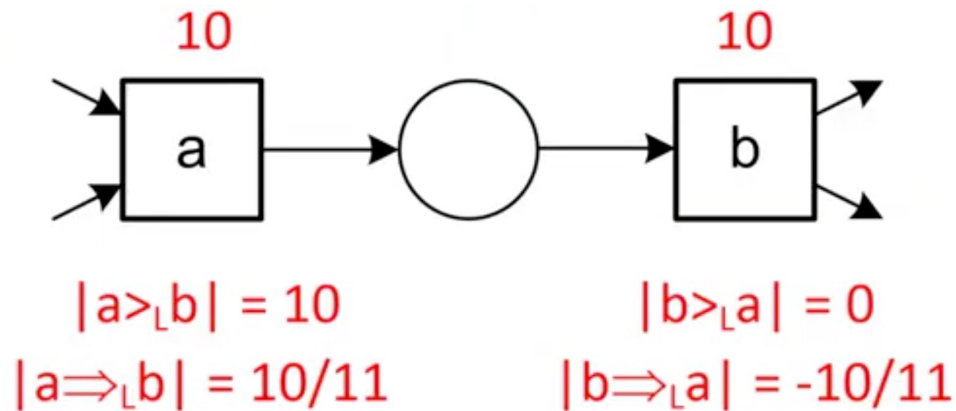
$$|a \Rightarrow_L b| = 10/11$$

$$|b >_L a| = 0$$

$$|b \Rightarrow_L a| = -10/11$$

Both measures need to be above a threshold to be included

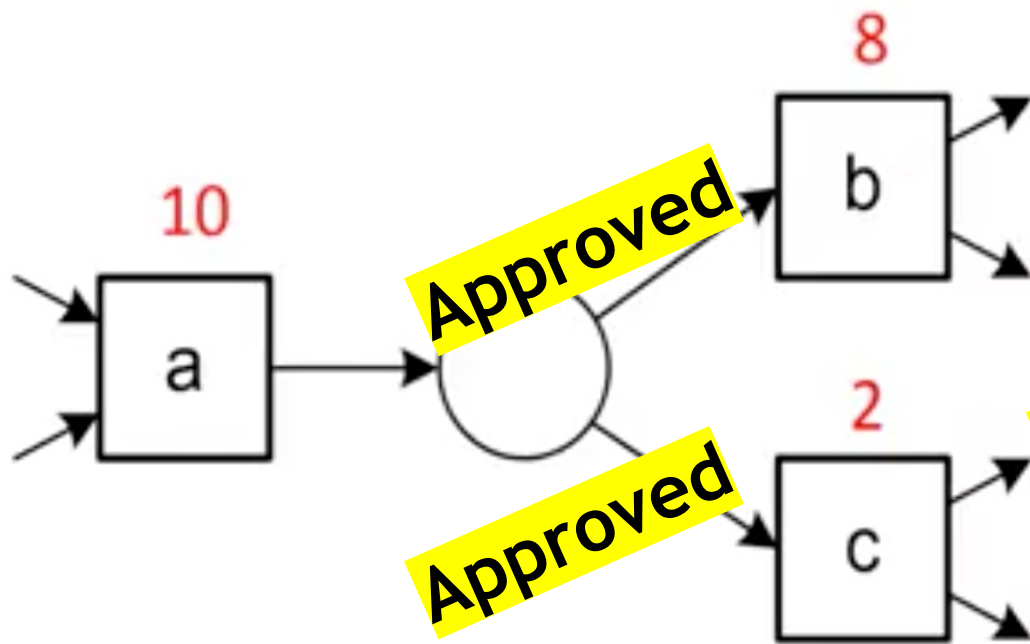
Included arc (assuming thresholds ≥ 1 and ≥ 0.5)



XOR-split pattern (assuming thresholds ≥ 1 and ≥ 0.5)

$$|a \succ_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a \succ_L b| - |b \succ_L a|}{|a \succ_L b| + |b \succ_L a| + 1} & \text{if } a \neq b \\ \frac{|a \succ_L a|}{|a \succ_L a| + 1} & \text{if } a = b \end{cases}$$



Approved

$$|a \succ_L b| = 8$$

$$|a \Rightarrow_L b| = 8/9$$

$$|b \succ_L a| = 0$$

$$|b \Rightarrow_L a| = -8/9$$

Approved

$$|a \succ_L c| = 2$$

$$|a \Rightarrow_L c| = 2/3$$

$$|b \succ_L c| = 0$$

$$|b \Rightarrow_L c| = 0/1$$

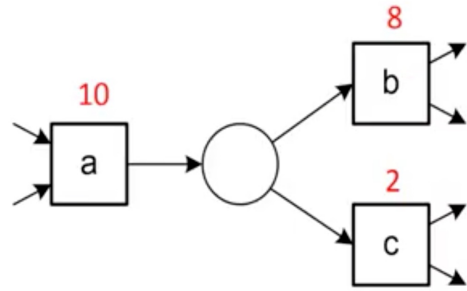
$$|c \succ_L a| = 0$$

$$|c \Rightarrow_L a| = -2/3$$

$$|c \succ_L b| = 0$$

$$|c \Rightarrow_L b| = 0/1$$

Included arcs (assuming thresholds ≥ 1 and ≥ 0.5)



$$|a \succ b| = 8$$

$$|a \Rightarrow b| = 8/9$$

$$|b \succ a| = 0$$

$$|b \Rightarrow a| = -8/9$$

$$|a \succ c| = 2$$

$$|a \Rightarrow c| = 2/3$$

$$|b \succ c| = 0$$

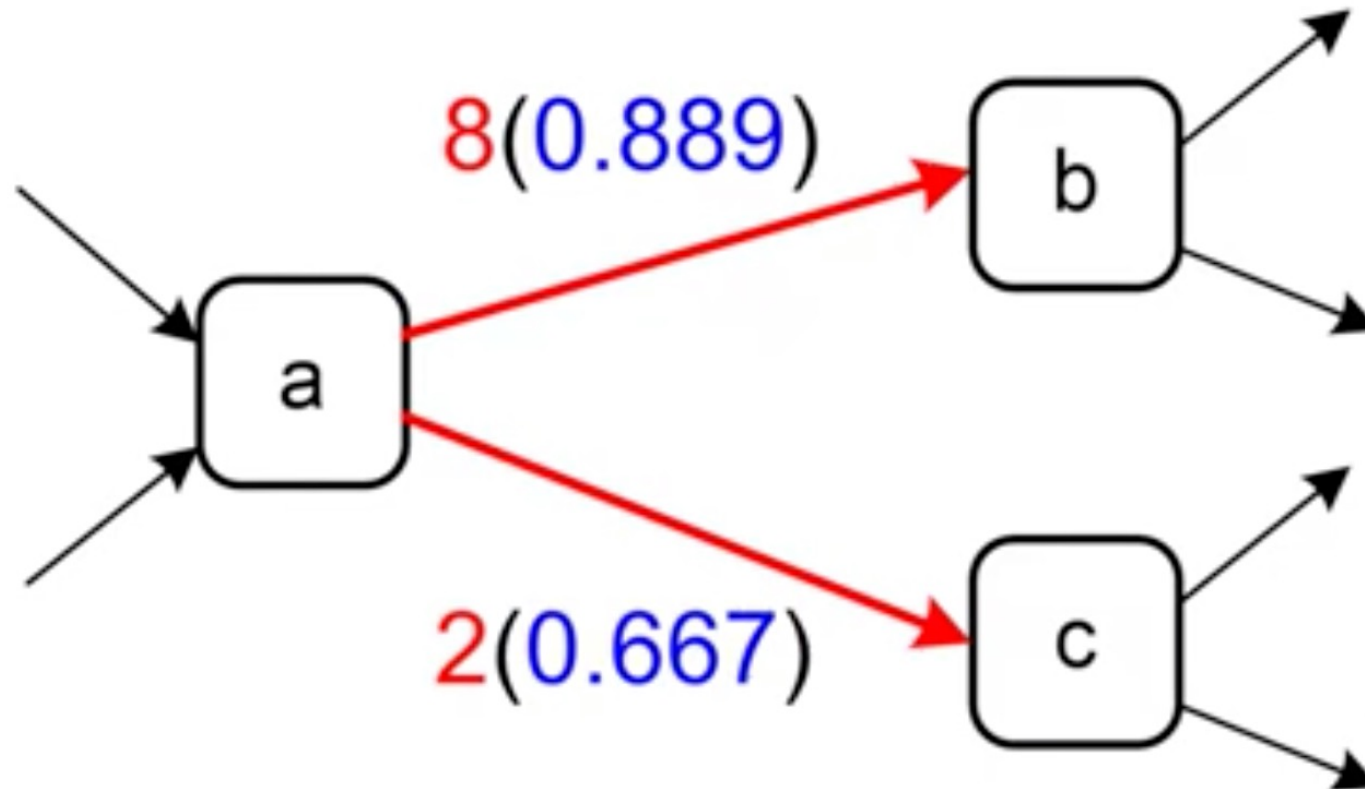
$$|b \Rightarrow c| = 0/1$$

$$|c \succ a| = 0$$

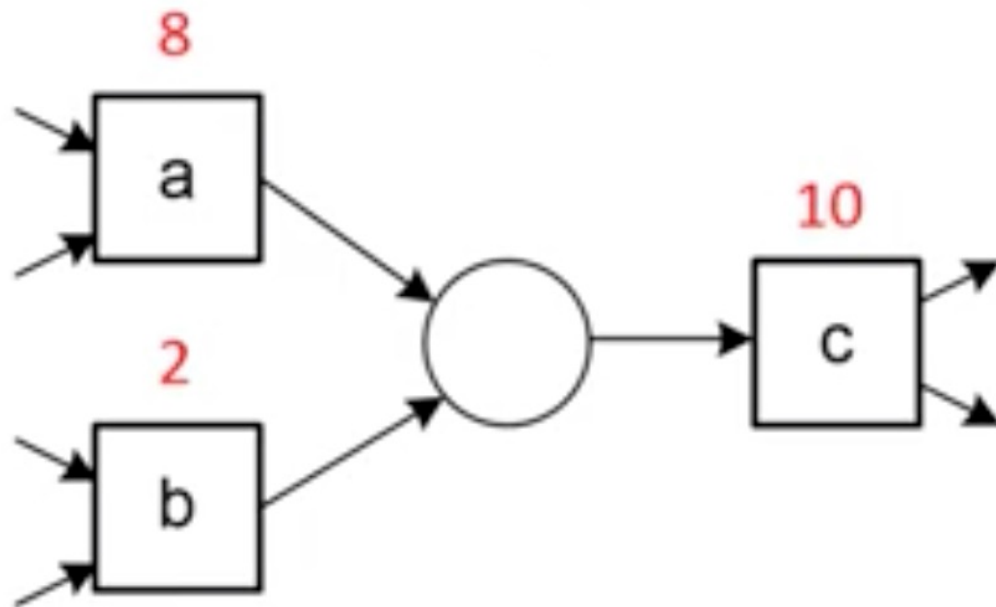
$$|c \Rightarrow a| = -2/3$$

$$|c \succ b| = 0$$

$$|c \Rightarrow b| = 0/1$$



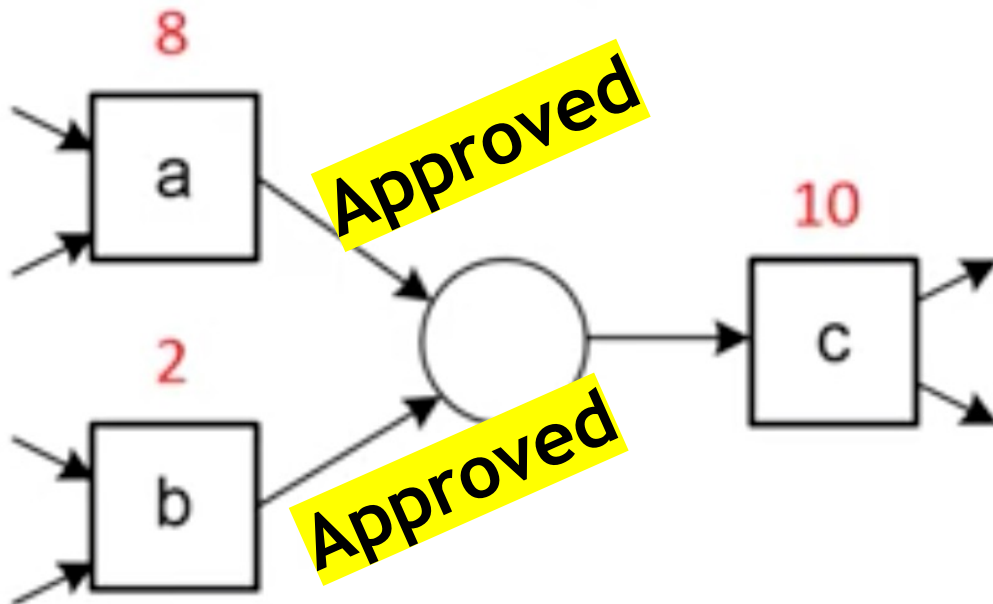
XOR-join pattern



XOR-join pattern

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$



$$|a >_L b| = 0$$

$$|a \Rightarrow_L b| = 0/1$$

$$|b >_L a| = 0$$

$$|b \Rightarrow_L a| = 0/1$$

Approved

$$|a >_L c| = 8$$

$$|a \Rightarrow_L c| = 8/9$$

Approved

$$|b >_L c| = 2$$

$$|b \Rightarrow_L c| = 2/3$$

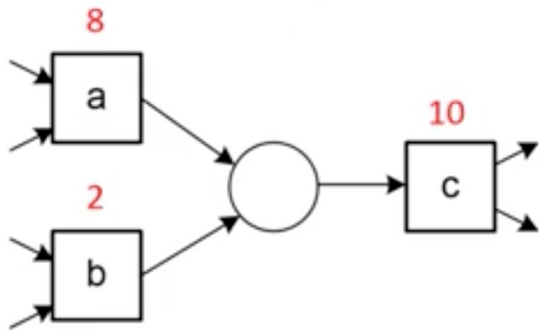
$$|c >_L a| = 0$$

$$|c \Rightarrow_L a| = -8/9$$

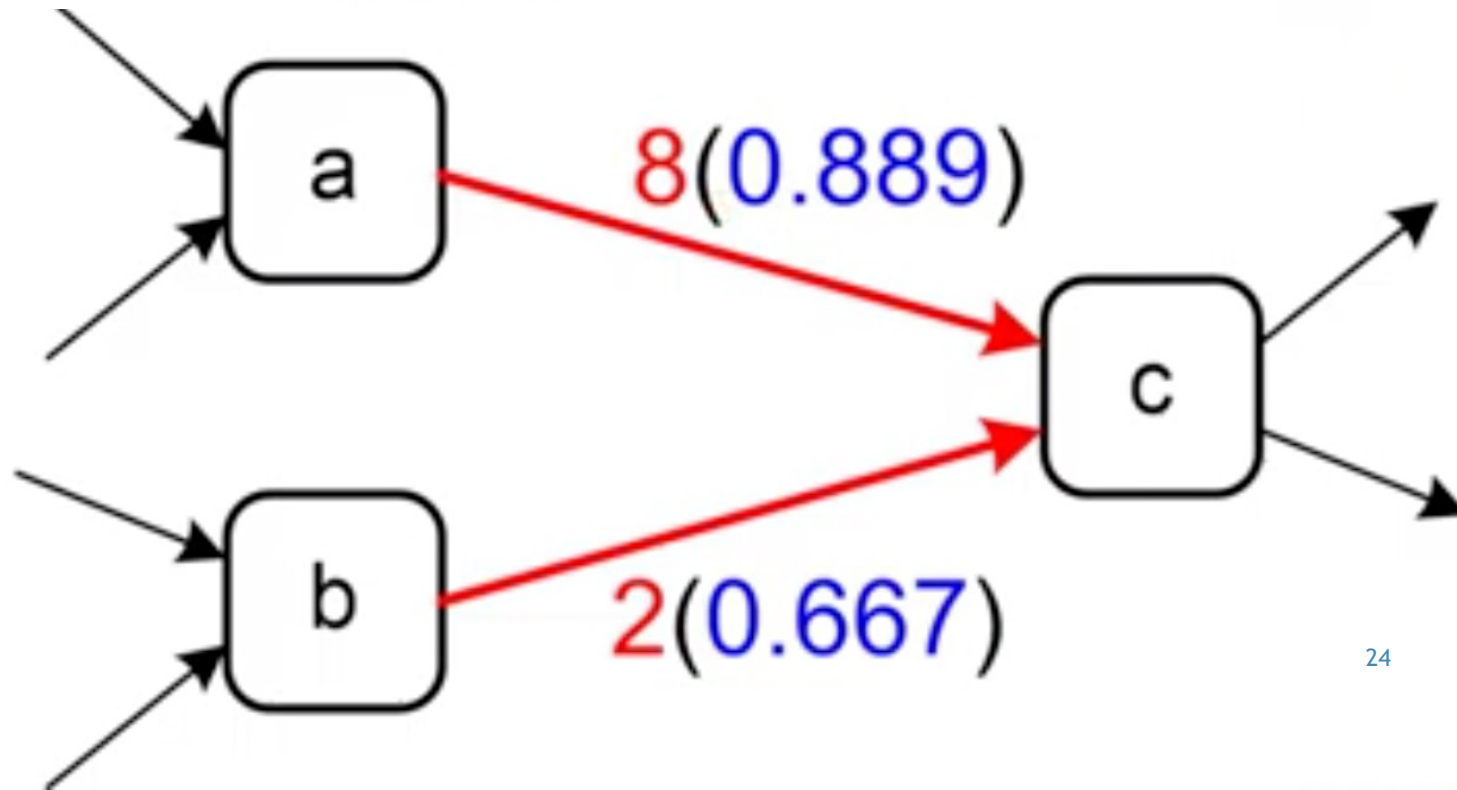
$$|c >_L b| = 0$$

$$|c \Rightarrow_L b| = -2/3$$

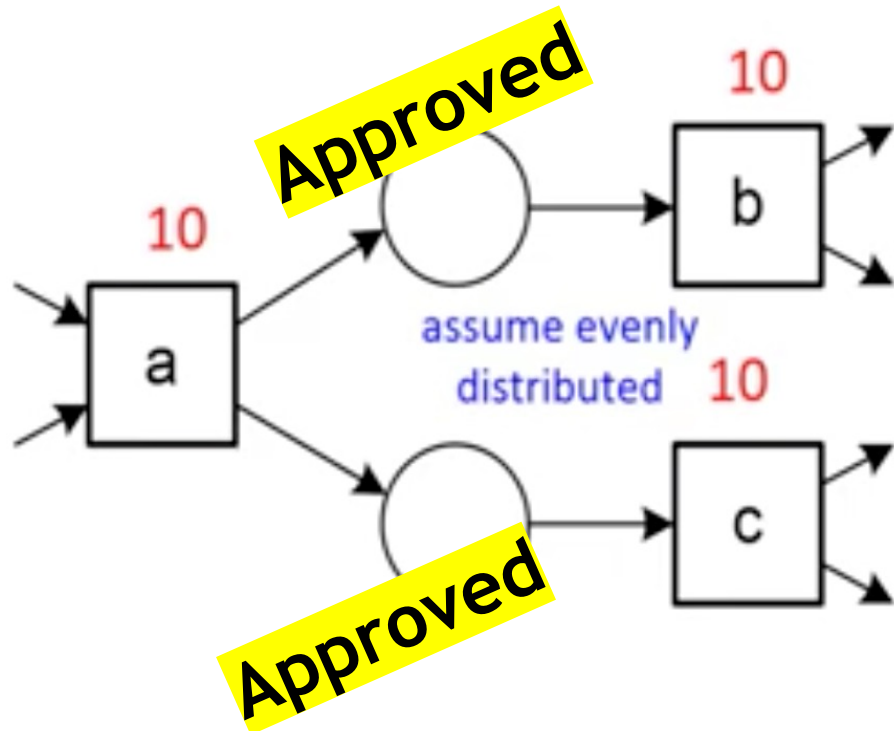
Included arcs (assuming thresholds ≥ 1)



$ a \succ b = 0$	$ b \succ a = 0$
$ a \Rightarrow b = 0/1$	$ b \Rightarrow a = 0/1$
$ a \succ c = 8$	$ b \succ c = 2$
$ a \Rightarrow c = 8/9$	$ b \Rightarrow c = 2/3$
$ c \succ a = 0$	$ c \succ b = 0$
$ c \Rightarrow a = -8/9$	$ c \Rightarrow b = -2/3$



AND-split pattern



Approved

Approved

$$|a>_l b| = 5$$

$$|a\Rightarrow_l b| = 5/6$$

$$|a>_l c| = 5$$

$$|a\Rightarrow_l c| = 5/6$$

$$|c>_l a| = 0$$

$$|c\Rightarrow_l a| = -5/6$$

$$|b>_l a| = 0$$

$$|b\Rightarrow_l a| = -5/6$$

$$|b>_l c| = 5$$

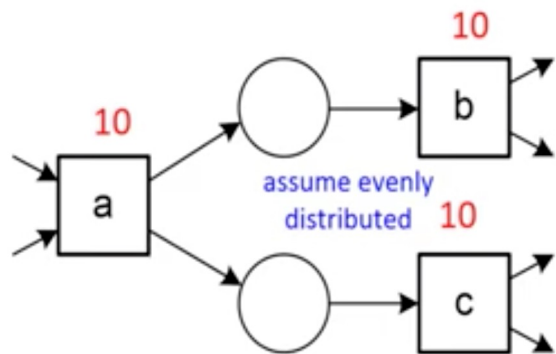
$$|b\Rightarrow_l c| = 0/11$$

$$|c>_l b| = 5$$

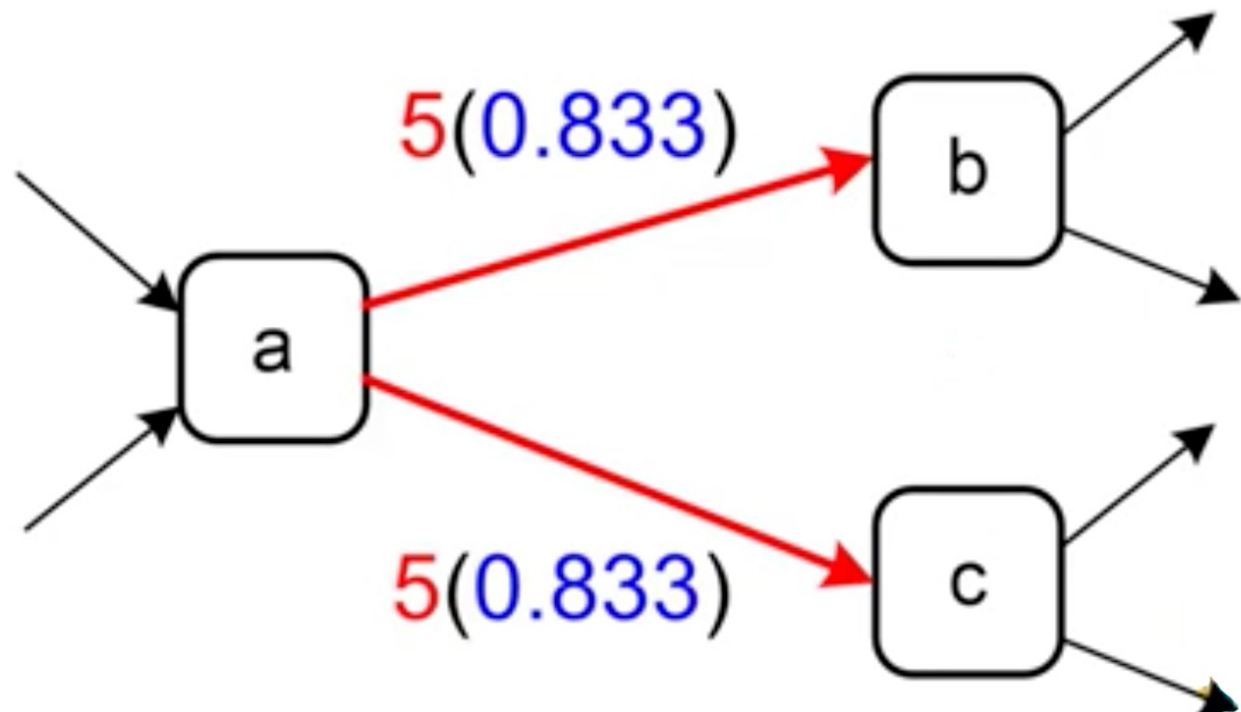
$$|c\Rightarrow_l b| = 0/11$$



Included arcs (assuming thresholds ≥ 1 and ≥ 0.5)



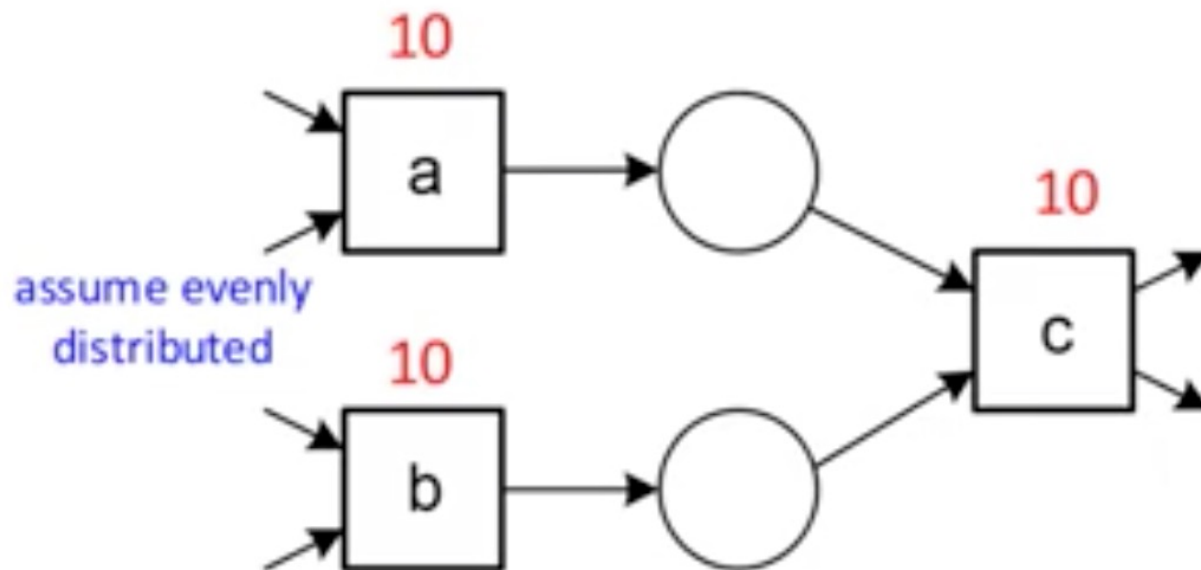
$ a >_l b = 5$	$ b >_l a = 0$
$ a \Rightarrow_l b = 5/6$	$ b \Rightarrow_l a = -5/6$
$ a >_l c = 5$	$ b >_l c = 5$
$ a \Rightarrow_l c = 5/6$	$ b \Rightarrow_l c = 0/11$
$ c >_l a = 0$	$ c >_l b = 5$
$ c \Rightarrow_l a = -5/6$	$ c \Rightarrow_l b = 0/11$



And-join pattern

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

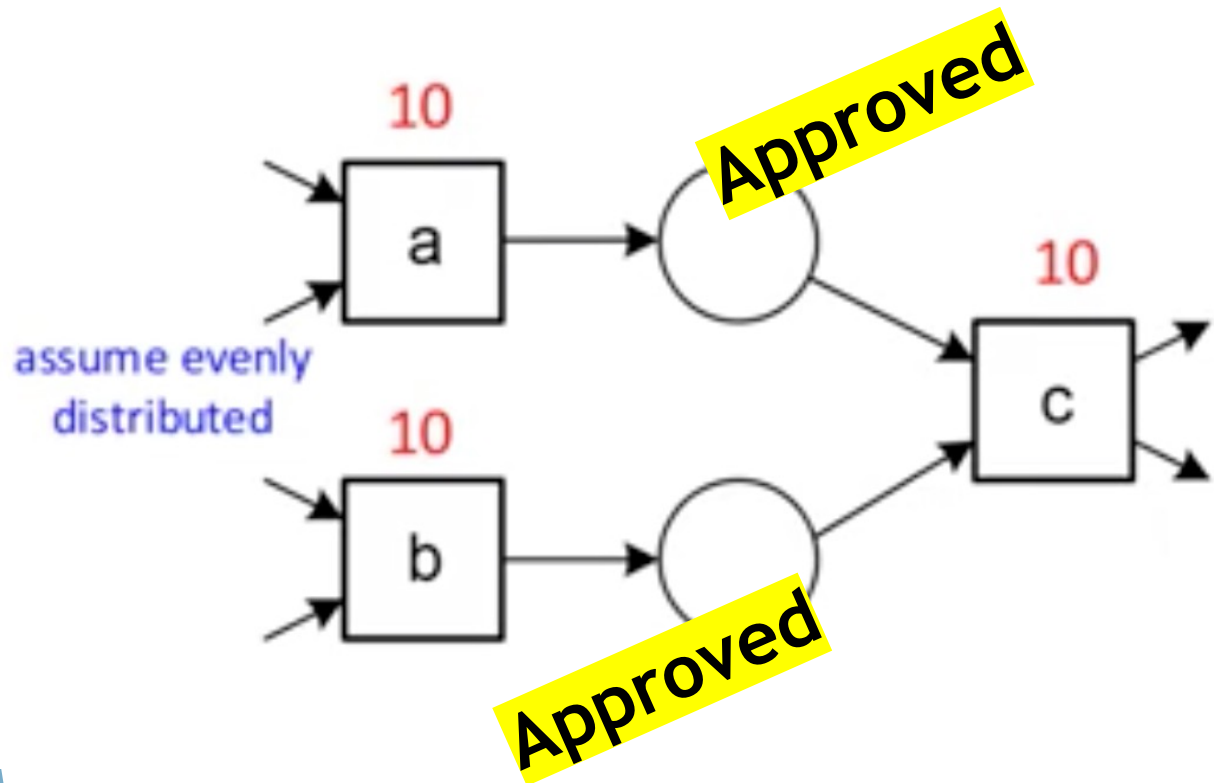
$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$



And-join pattern

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$



$$|a >_L b| = 5$$

$$|a \Rightarrow_L b| = 0/11$$

!

$$|b >_L a| = 5$$

$$|b \Rightarrow_L a| = 0/11$$

$$|a >_L c| = 5$$

$$|a \Rightarrow_L c| = 5/6$$

$$|b >_L c| = 5$$

$$|b \Rightarrow_L c| = 5/6$$

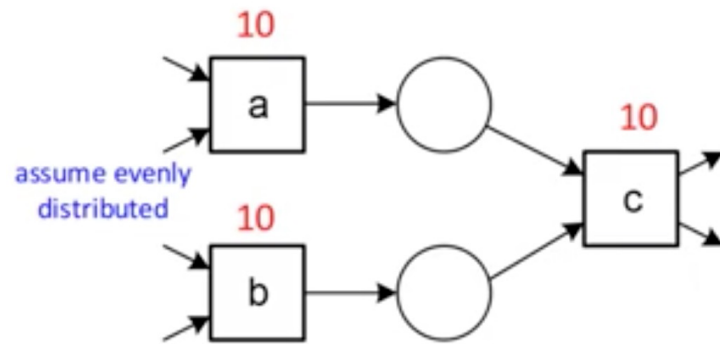
$$|c >_L a| = 0$$

$$|c \Rightarrow_L a| = -5/6$$

$$|c >_L b| = 0$$

$$|c \Rightarrow_L b| = -5/6$$

Included arcs (assuming thresholds ≥ 1 and ≥ 0.5)



$$|a \succ_L b| = 5$$

$$|a \Rightarrow_L b| = 0/11$$

$$|b \succ_L a| = 5$$

$$|b \Rightarrow_L a| = 0/11$$

$$|a \succ_L c| = 5$$

$$|a \Rightarrow_L c| = 5/6$$

$$|b \succ_L c| = 5$$

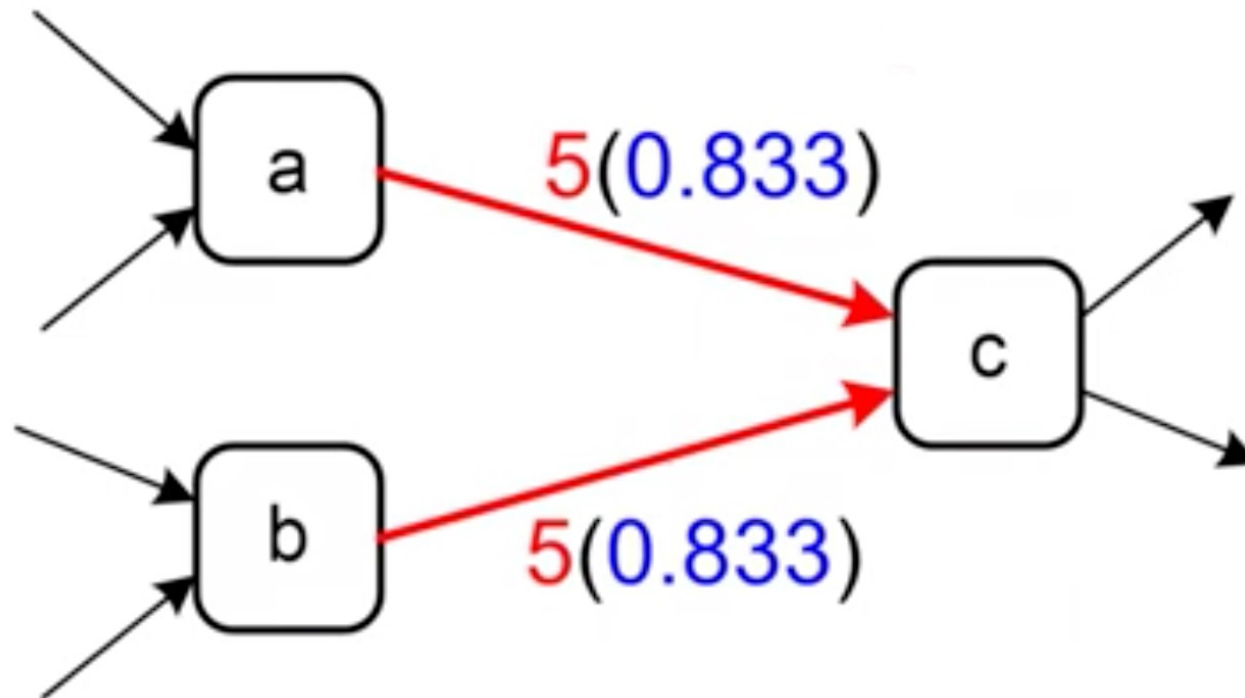
$$|b \Rightarrow_L c| = 5/6$$

$$|c \succ_L a| = 0$$

$$|c \Rightarrow_L a| = -5/6$$

$$|c \succ_L b| = 0$$

$$|c \Rightarrow_L b| = -5/6$$

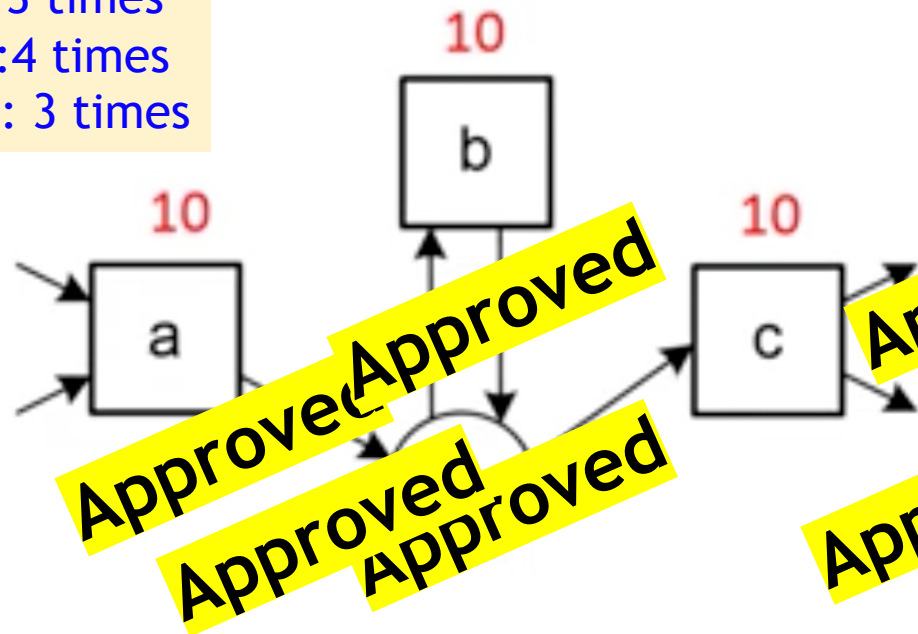


Loop pattern

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$

Assume
ac: 3 times
abc: 4 times
abbc: 3 times



$$|a >_L a| = 0$$

$$|a \Rightarrow_L a| = 0/1$$

$$|b >_L a| = 0$$

$$|b \Rightarrow_L a| = -7/8$$

$$|c >_L a| = 0$$

$$|c \Rightarrow_L a| = -3/4$$

$$|a >_L b| = 7$$

$$|a \Rightarrow_L b| = 7/8$$

$$|b >_L b| = 3$$

$$|b \Rightarrow_L b| = 3/4$$

$$|c >_L b| = 0$$

$$|c \Rightarrow_L b| = -7/8$$

$$|a >_L c| = 3$$

$$|a \Rightarrow_L c| = 3/4$$

$$|b >_L c| = 7$$

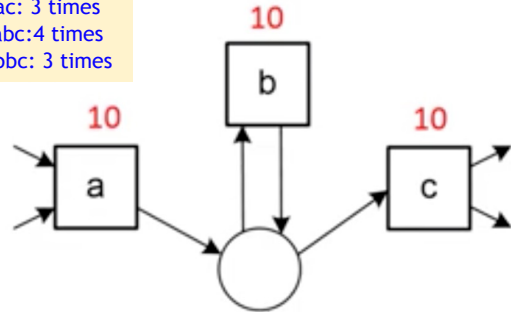
$$|b \Rightarrow_L c| = 7/8$$

$$|c >_L c| = 0$$

$$|c \Rightarrow_L c| = 0/1$$

Included arcs (assuming thresholds ≥ 1 and ≥ 0.5)

Assume
ac: 3 times
abc: 4 times
abbc: 3 times



$$|a \triangleright_L a| = 0$$

$$|a \Rightarrow_L a| = 0/1$$

$$|b \triangleright_L a| = 0$$

$$|b \Rightarrow_L a| = -7/8$$

$$|c \triangleright_L a| = 0$$

$$|c \Rightarrow_L a| = -3/4$$

$$|a \triangleright_L b| = 7$$

$$|a \Rightarrow_L b| = 7/8$$

$$|b \triangleright_L b| = 3$$

$$|b \Rightarrow_L b| = 3/4$$

$$|c \triangleright_L b| = 0$$

$$|c \Rightarrow_L b| = -7/8$$

$$|a \triangleright_L c| = 3$$

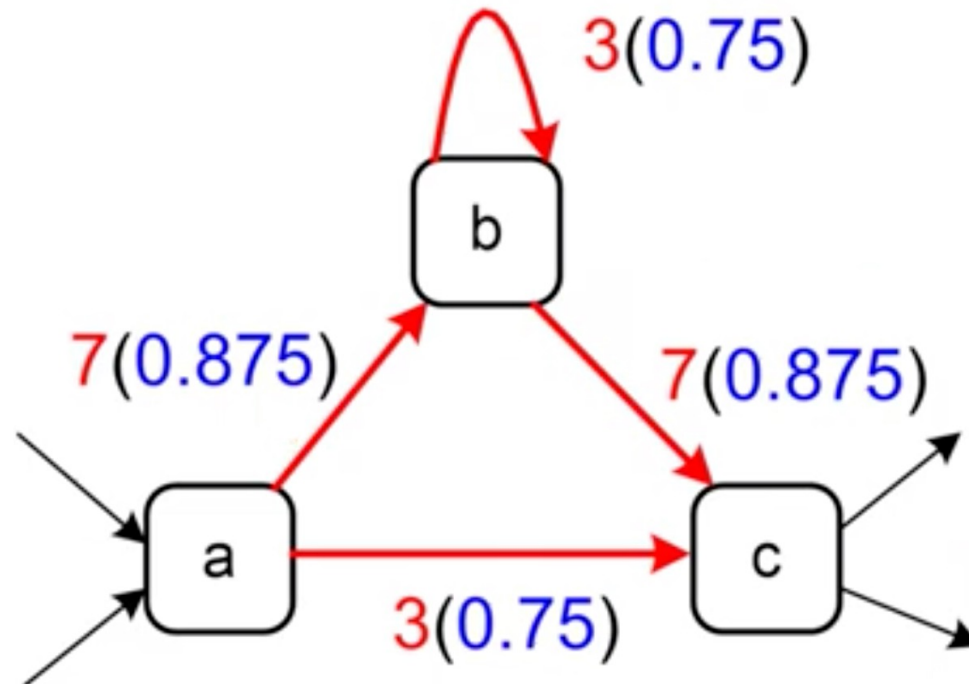
$$|a \Rightarrow_L c| = 3/4$$

$$|b \triangleright_L c| = 7$$

$$|b \Rightarrow_L c| = 7/8$$

$$|c \triangleright_L c| = 0$$

$$|c \Rightarrow_L c| = 0/1$$



Home Work

- Compute the dependency measures: $|a \Rightarrow_L b|$ and $|d \Rightarrow_L d|$ for the given event log.

$L = [\langle a, e \rangle^5, \langle a, b, c, e \rangle^{10}, \langle a, c, b, e \rangle^{10}, \langle a, b, e \rangle^1, \langle a, c, e \rangle^1, \langle a, d, e \rangle^{10}, \langle a, d, d, e \rangle^2, \langle a, d, d, d, e \rangle^1]$

$$|a >_L b| = \sum_{\sigma \in L} L(\sigma) \times |\{1 \leq i < |\sigma| \mid \sigma(i) = a \wedge \sigma(i+1) = b\}|$$

$$|a \Rightarrow_L b| = \begin{cases} \frac{|a >_L b| - |b >_L a|}{|a >_L b| + |b >_L a| + 1} & \text{if } a \neq b \\ \frac{|a >_L a|}{|a >_L a| + 1} & \text{if } a = b \end{cases}$$

Reading Material

- ▶ Chapter 3 & 7: Aalst