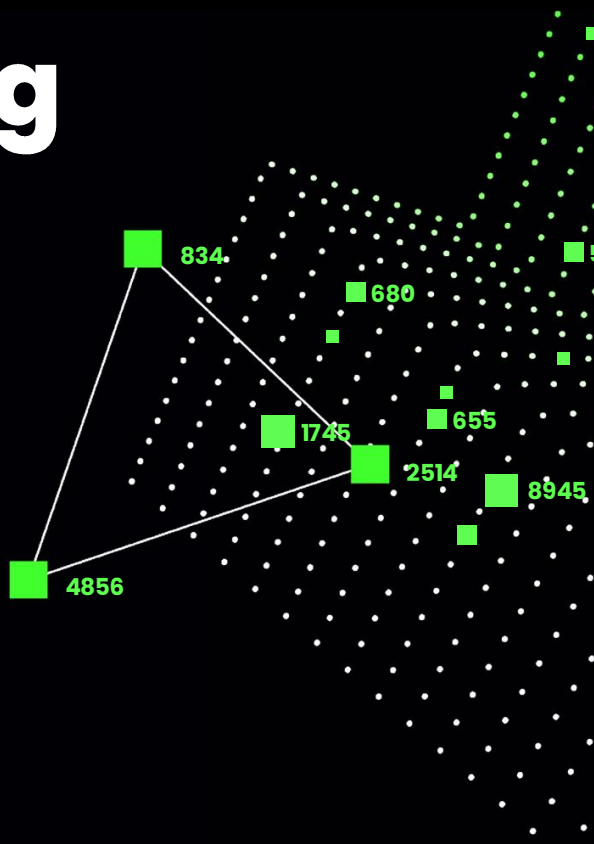
celonis

Chair of Process and Data Science | RWTH AACHEN UNIVERSITY

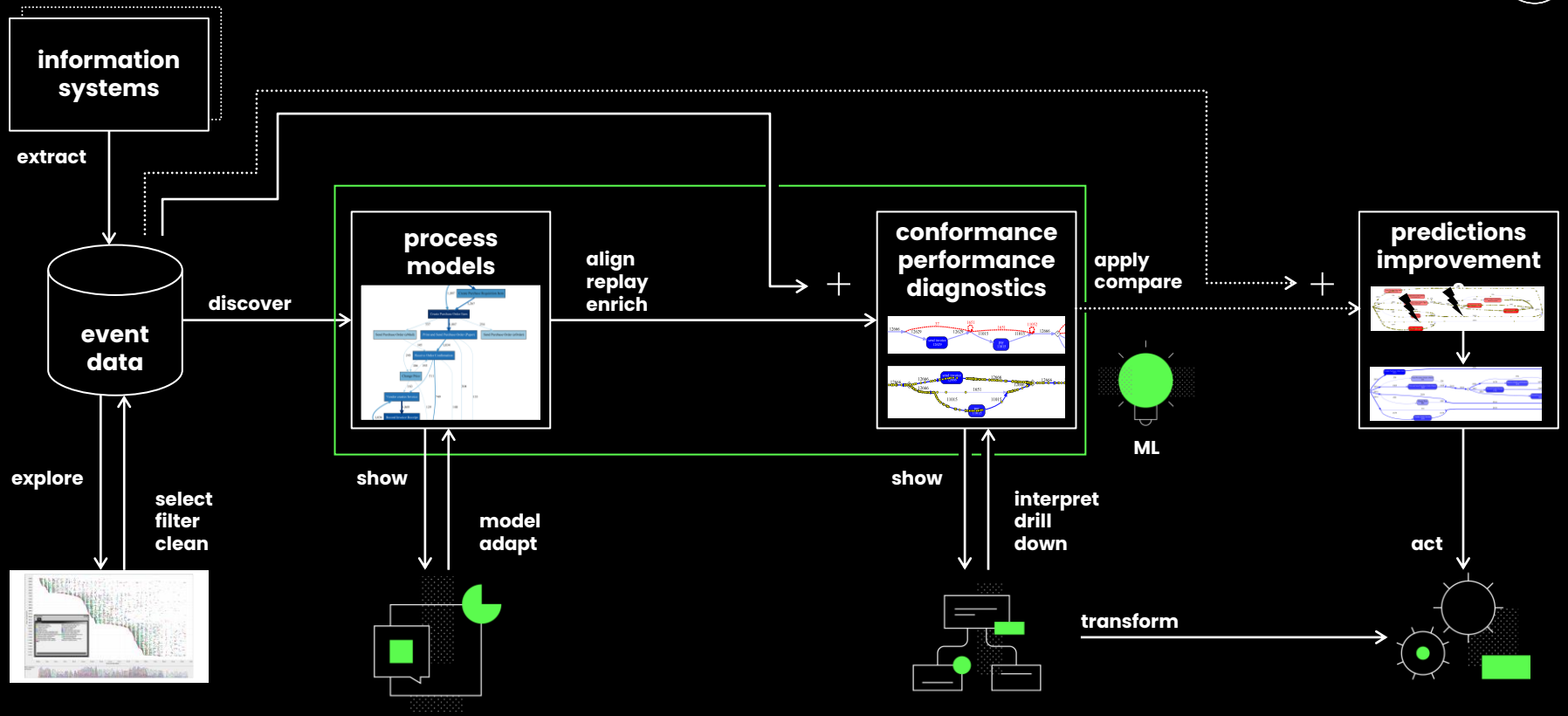
Conformance Checking

Checking compliance and analyzing deviations

Process Mining: From Theory to Execution



prof.dr.ir. Wil van der Aalst
www.vdaalst.com @wvdaalst | www.pads.rwth-aachen.de



© Wil van der Aalst (use only with permission & acknowledgements)

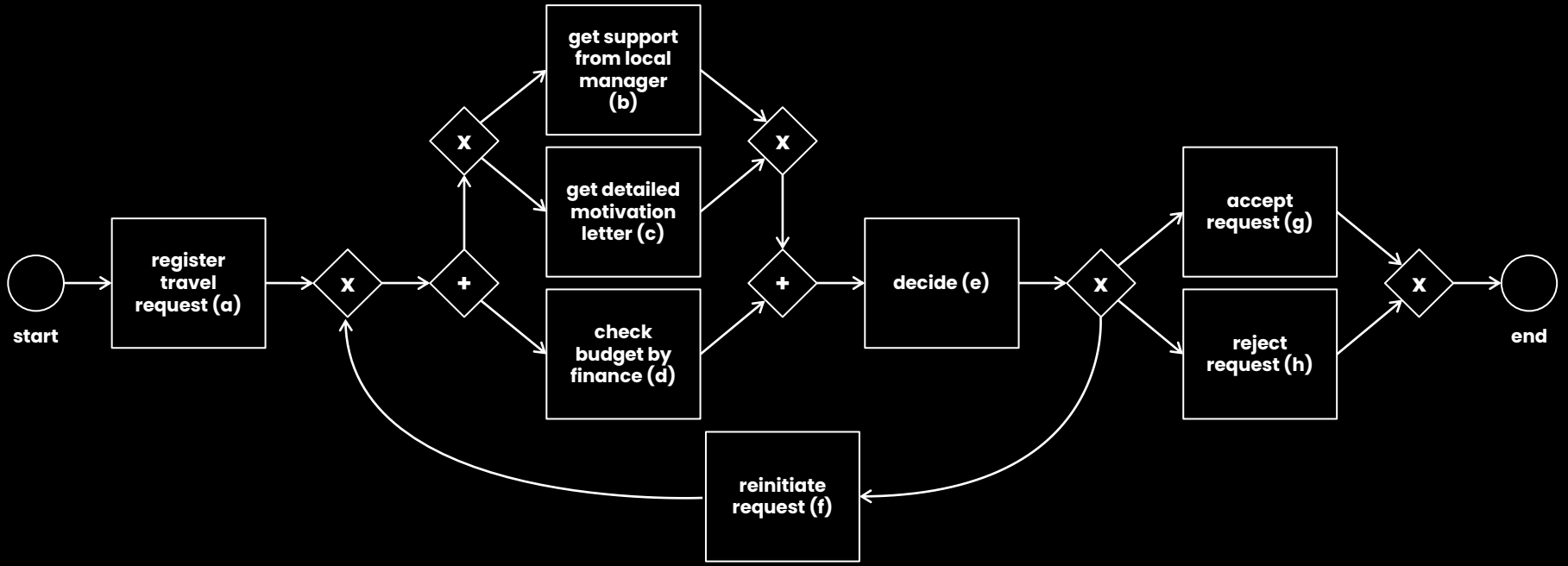
Input: Event data and process model

conforming

a	b	d	e	g

deviating

a	b	c	e	h

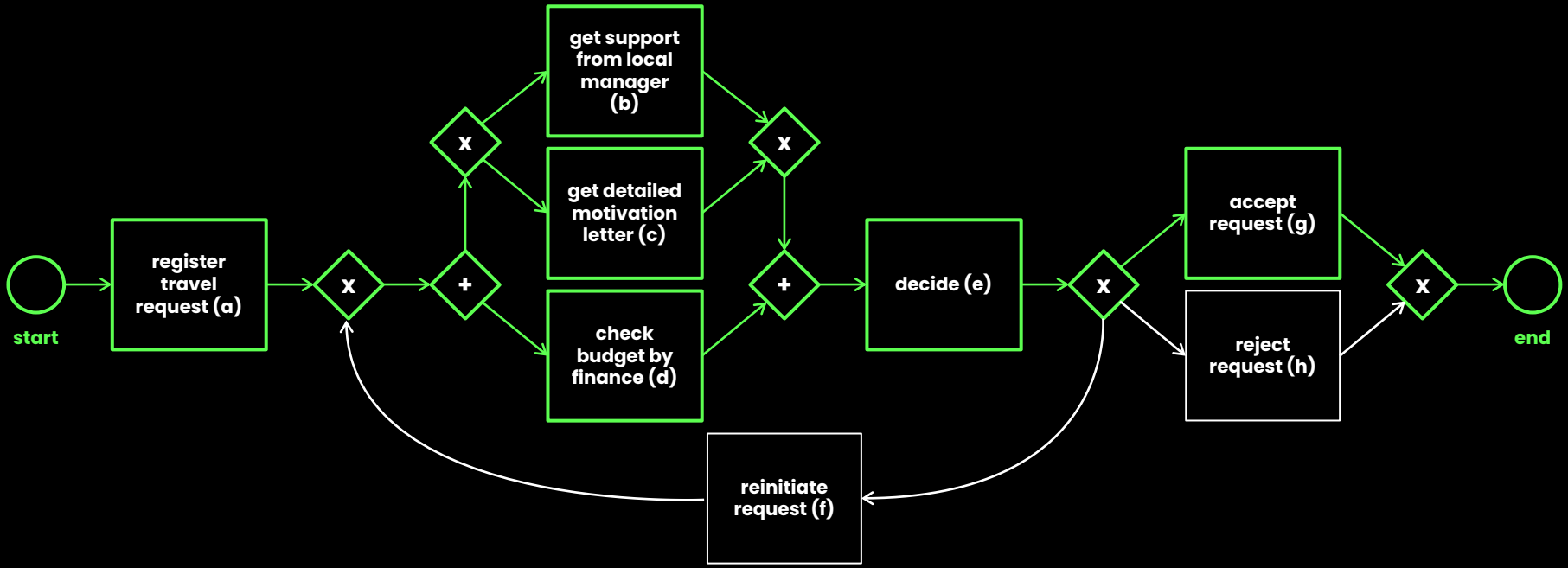


© Wil van der Aalst (use only with permission & acknowledgements)

Conforming case

conforming

a	b	d	e	g

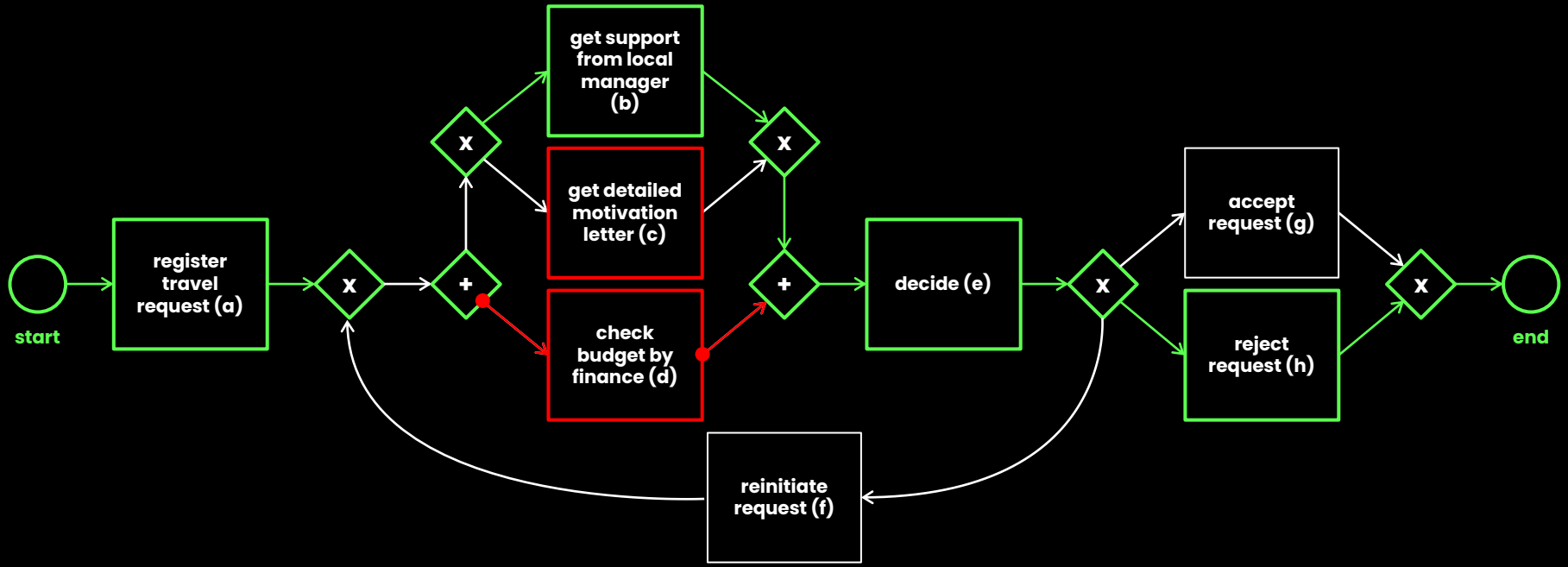


© Wil van der Aalst (use only with permission & acknowledgements)

Deviating case

deviating

a b ~~c~~ d h



© Wil van der Aalst (use only with permission & acknowledgements)

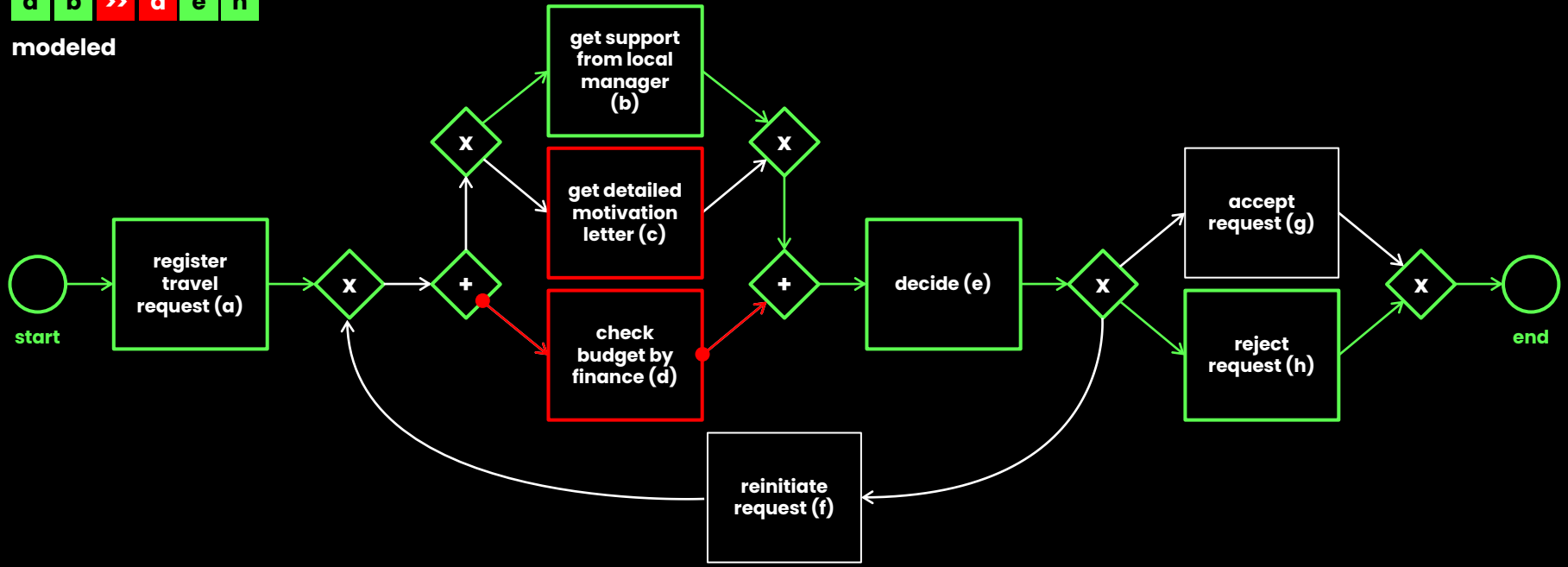
Alignment: Relating observed and modeled behavior

observed

a	b	c	>>	e	h

a	b	>>	d	e	h

modeled



© Wil van der Aalst (use only with permission & acknowledgements)

Alignments do not need to be unique

1

a	b	c	»	e	h
a	b	»	d	e	h

2

a	b	c	»	e	h
a	»	c	d	e	h

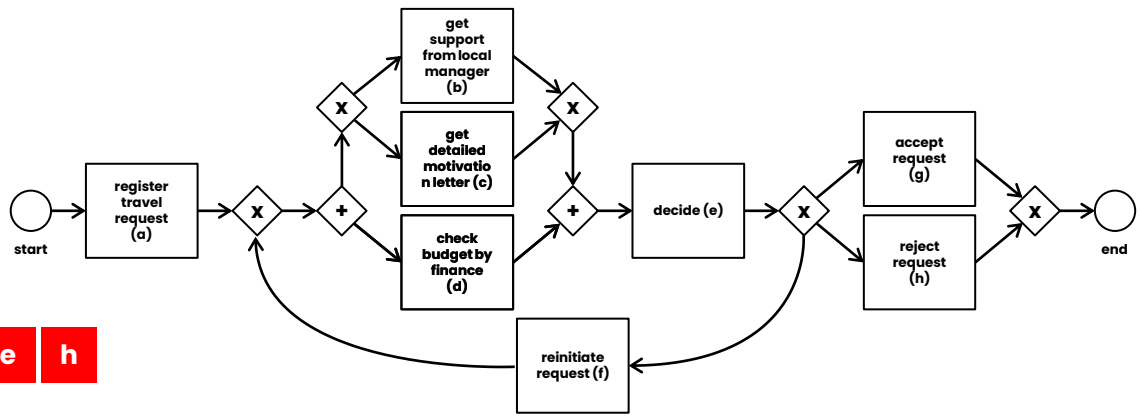
3

a	b	»	c	e	h
a	b	d	»	e	h

4

a	»	b	c	e	h
a	d	»	c	e	h

+ 2 more

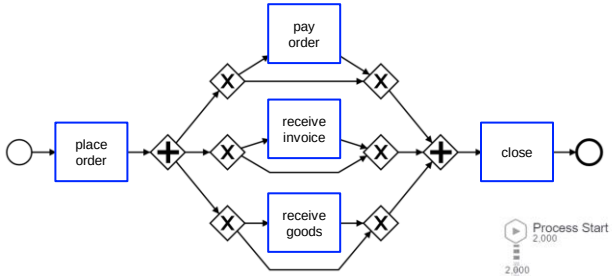


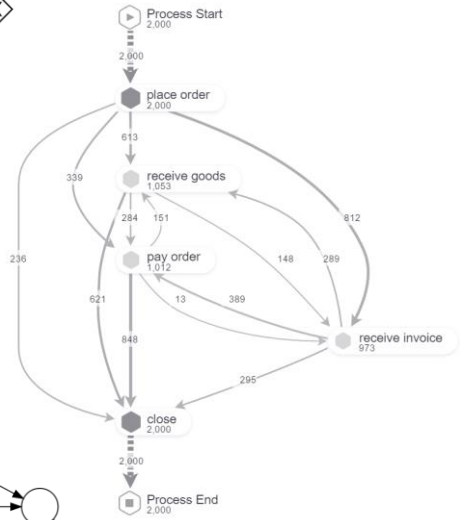
a	b	c	e	h

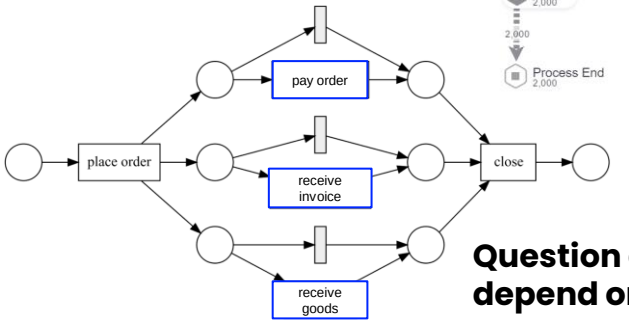
© Wil van der Aalst (use only with permission & acknowledgements)

Find a “nearest path” in the model

285 traces 14.25%	place order	receive goods	close	
251 traces 12.55%	place order	pay order	close	
238 traces 11.90%	place order	receive invoice	pay order	close
236 traces 11.80%	place order	close		
216 traces 10.80%	place order	receive goods	close	
183 traces 9.15%	place order	receive invoice	receive goods	close
177 traces 8.85%	place order	receive goods	pay order	close
103 traces 5.15%	place order	receive invoice	receive goods	pay order	close
79 traces 3.95%	place order	receive goods	receive invoice	pay order	close
78 traces 3.90%	place order	pay order	receive goods	close
72 traces 3.60%	place order	receive invoice	pay order	receive goods	close
68 traces 3.40%	place order	receive goods	receive invoice	close
6 traces 0.30%	place order	pay order	receive invoice	close
4 traces 0.20%	place order	receive goods	pay order	receive invoice	close
3 traces 0.15%	place order	pay order	receive invoice	receive goods	close
1 trace 0.05%	place order	pay order	receive goods	receive invoice	close







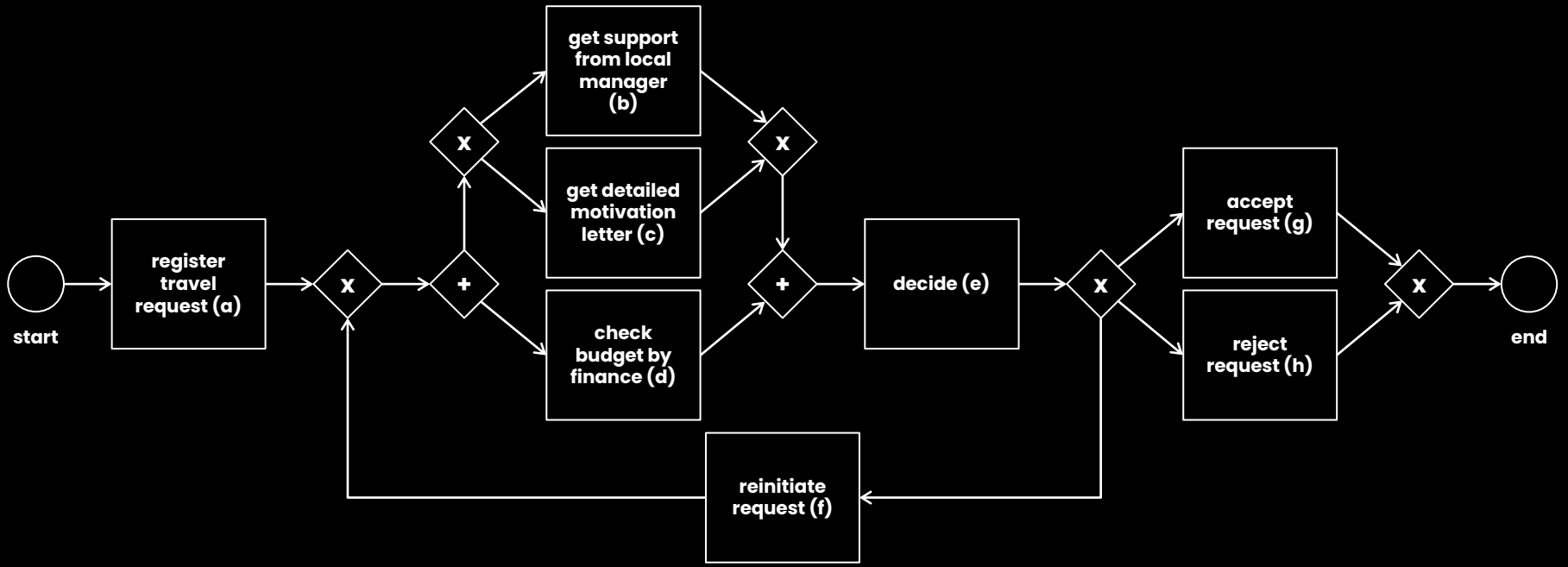
Question does not depend on notation!

© Wil van der Aalst (use only with permission & acknowledgements)

Provide an optimal alignment

observed

a	b	d	d	e	f	f	b	e	g



© Wil van der Aalst (use only with permission & acknowledgements)

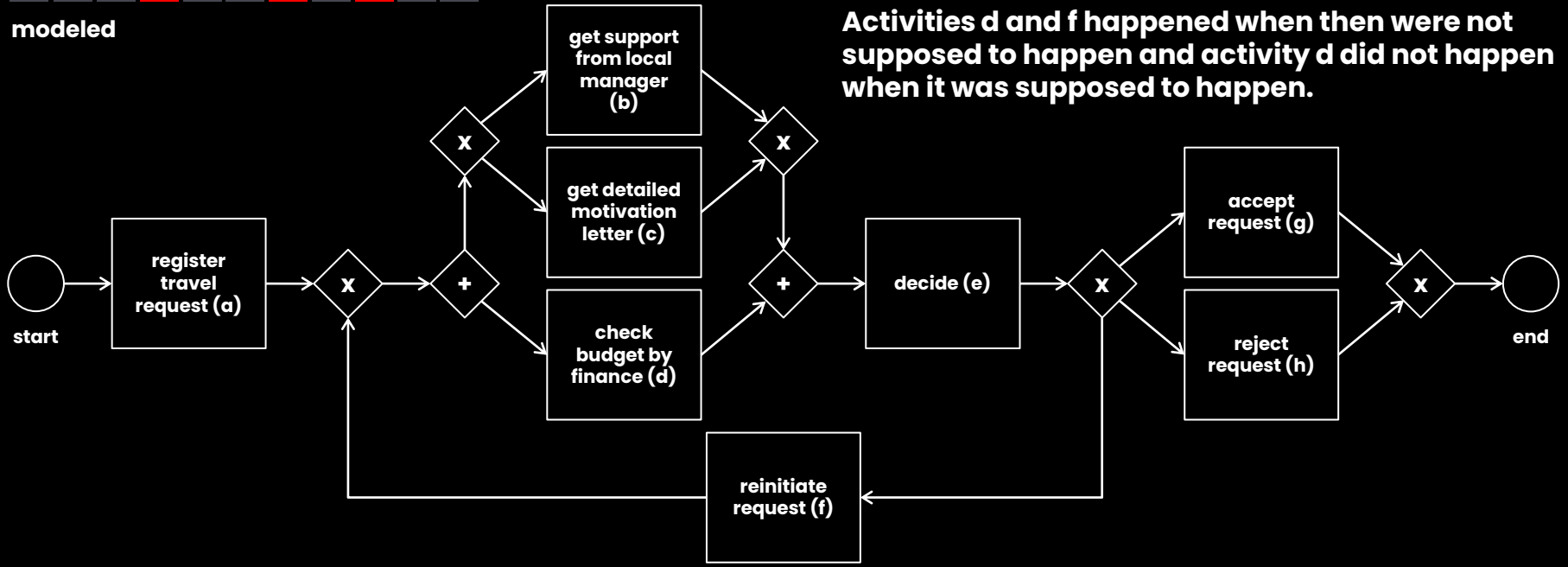
Provide an optimal alignment

observed

a	b	d	d	e	f	f	b	>>	e	g

a	b	d	>>	e	f	>>	b	d	e	g

modeled

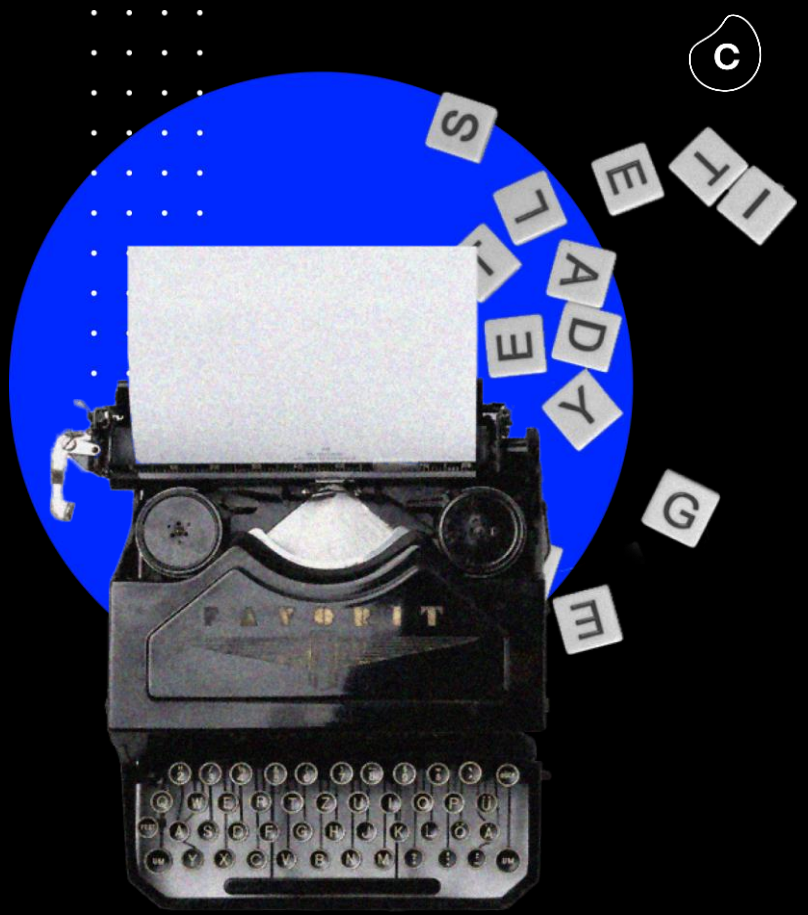


Activities d and f happened when then were not supposed to happen and activity d did not happen when it was supposed to happen.

© Wil van der Aalst (use only with permission & acknowledgements)

Conformance checking questions (1/2)

- A process model describes a **set of possible traces**
- An event log contains a **multiset of observed traces**



Basic questions:

- Which traces in the log are “fitting” the model?
- What is the percentage of “fitting” traces?
- Which modeled activities are often skipped in reality?
- Which observed activities are often impossible according to the process model?

© Wil van der Aalst (use only with permission & acknowledgements)

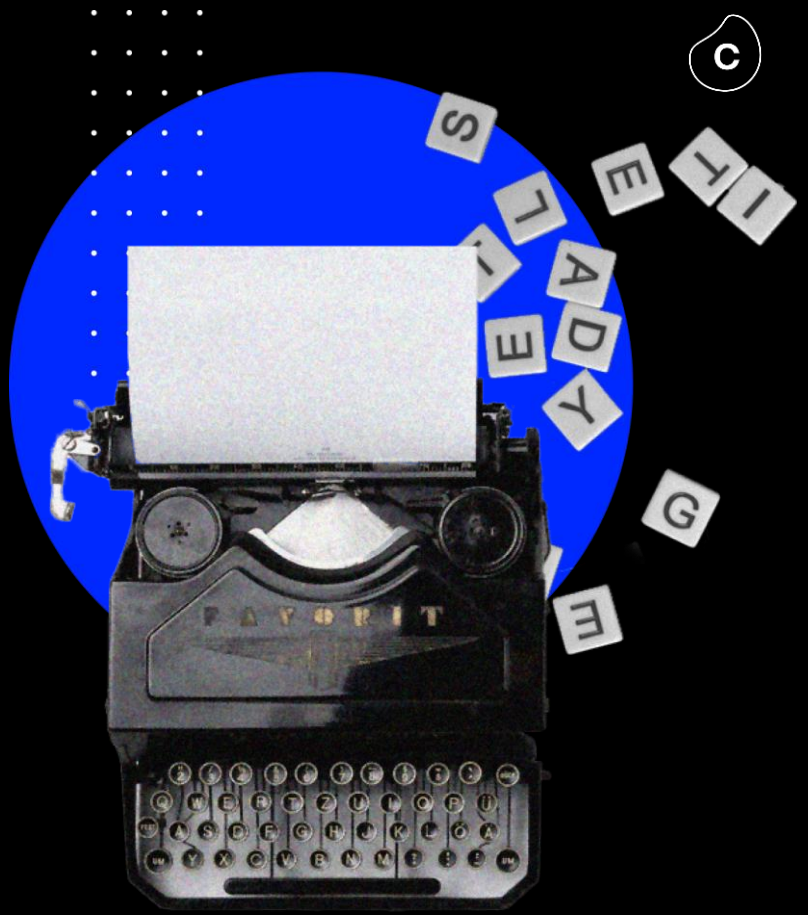
Conformance checking questions (2/2)

Advanced questions

- Which (shorter) traces, possible according to the model, never happen?
- What is the precision of the model (fraction of modeled behavior actually observed)?
- What do specific deviations have in common?

There are many ways to quantify conformance:

- Trace or event level
- Fitness (recall)/precision
- Adding time, resources, etc.



© Wil van der Aalst (use only with permission & acknowledgements)

Computing alignments

210 traces 7.91%	create purchase	create purchase order	print and send purchase	receive order	record goods receipt	vendor creates invoice	record invoice receipt
170 traces 5.41%	create purchase order	print and send purchase	receive order	record goods receipt	vendor creates invoice	record invoice receipt	
109 traces 4.11 %	create purchase order	print and send purchase					
32 traces 3.09%	create purchase order	print and send purchase	receive order	record goods receipt			
78 traces 2.94%	create purchase	create purchase order	print and send purchase	receive order	vendor creates invoice	record invoice receipt	record goods receipt
70 traces 2.64%	create purchase order	print and send purchase	receive order	vendor creates invoice	record invoice receipt	record goods receipt	
52 traces 2.34%	create purchase order	print and send purchase	receive order				
57 traces 2.15%	create purchase order	print and send purchase	receive order	record goods receipt			
54 traces 2.03%	create purchase order	print and send purchase	record goods receipt	vendor creates invoice	record invoice receipt		
48 traces 1.81%	create purchase	create purchase order	print and send purchase	record goods receipt	vendor creates invoice	record invoice receipt	
43 traces 1.62%	create purchase order	print and send purchase	record goods receipt				

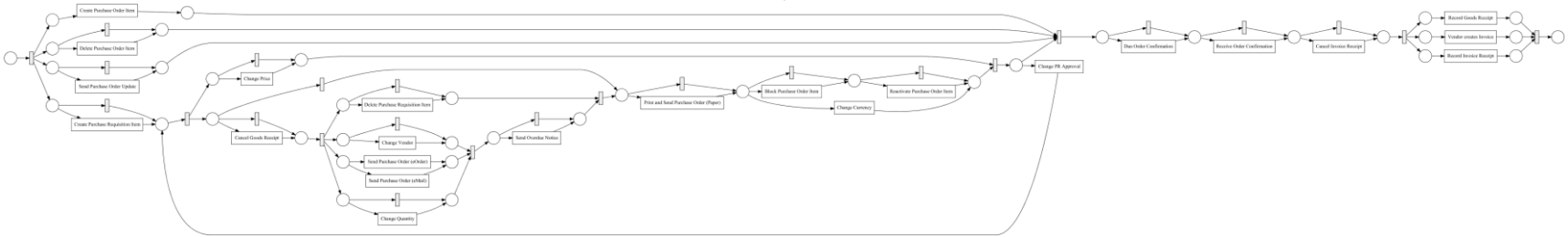
+650 more variants

The **gold** standard.

Needed for **advanced analysis** (e.g., predictions, performance analysis, and root cause analysis).

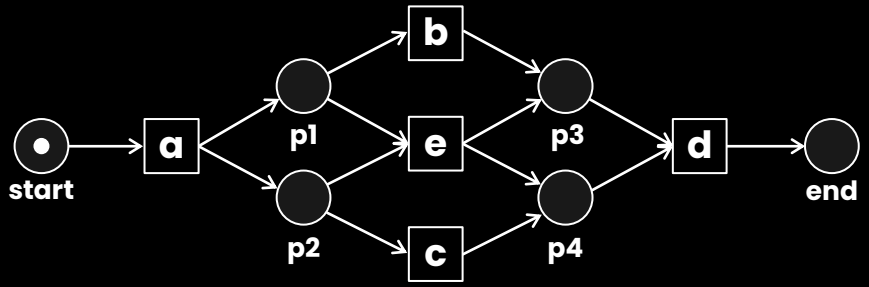
Computationally **challenging**.

Many optimal **alignments**.



© Wil van der Aalst (use only with permission & acknowledgements)

Alternative 1: Comparing footprints



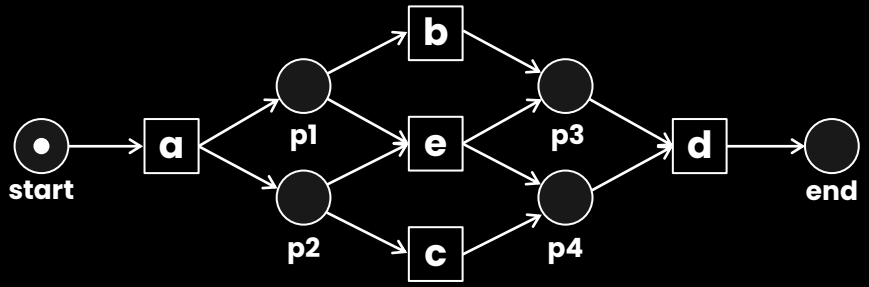
a	b	c	d
a	c	b	d
a	b	c	e	d

Footprint	a	b	c	d	e
a	#	→	→	#	→
b	←	#	‖	→	#
c	←	‖	#	→	#
d	#	←	←	#	←
e	←	#	#	→	#

Footprint	a	b	c	d	e
a	#	→	→	#	#
b	←	#	‖	→	#
c	←	‖	#	→	→
d	#	←	←	#	←
e	#	#	←	→	#

© Wil van der Aalst (use only with permission & acknowledgements)

Alternative 1: Comparing footprints



a	b	c	d
a	c	b	d
a	b	c	e	d

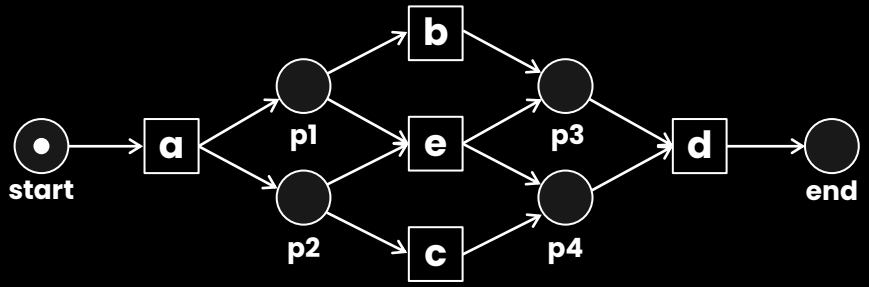
$$\frac{25-4}{25} = 0.84$$

Footprint	a	b	c	d	e
a	#	→	→	#	→
b	←	#	‖	→	#
c	←	‖	#	→	#
d	#	←	←	#	←
e	←	#	#	→	#

Footprint	a	b	c	d	e
a	#	→	→	#	#
b	←	#	‖	→	#
c	←	‖	#	→	→
d	#	←	←	#	←
e	#	#	←	→	#

© Wil van der Aalst (use only with permission & acknowledgements)

Alternative 2: Token based replay



a	b	c	d	
a	c	b	d	
a	b	c	e	d

p — **produced tokens**

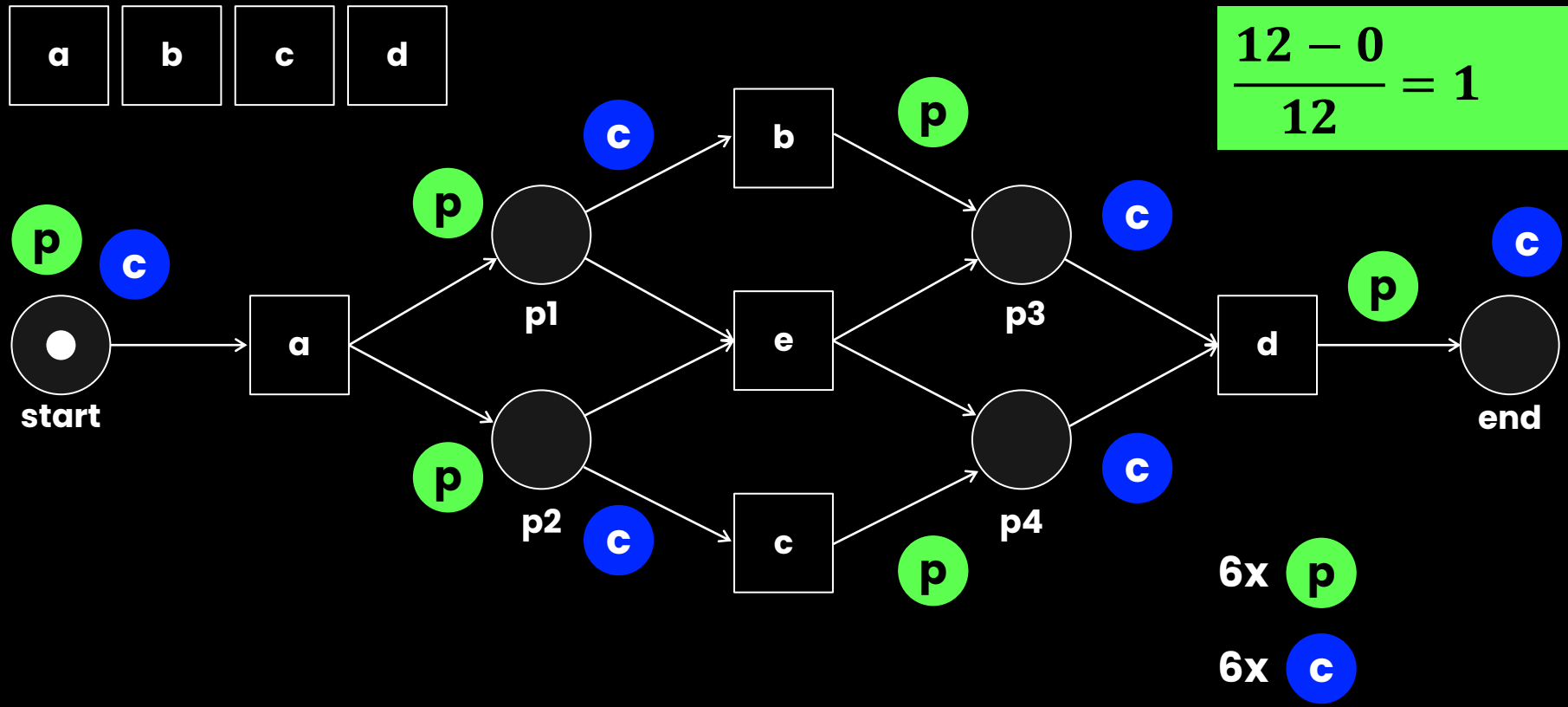
c — **consumed tokens**

m — **missing tokens** (something could not happen)

r — **remaining tokens** (something should have happened)

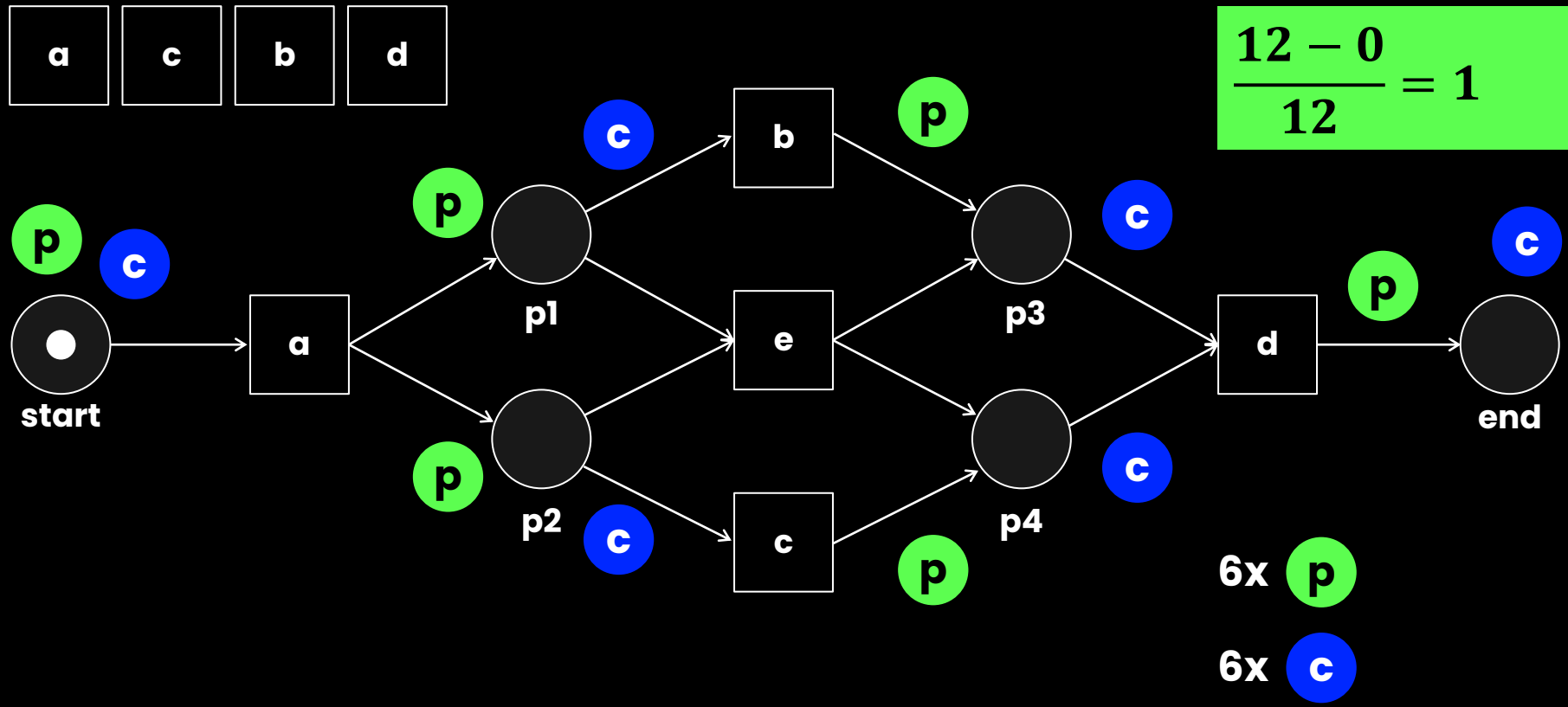
© Wil van der Aalst (use only with permission & acknowledgements)

Alternative 2: Token based replay



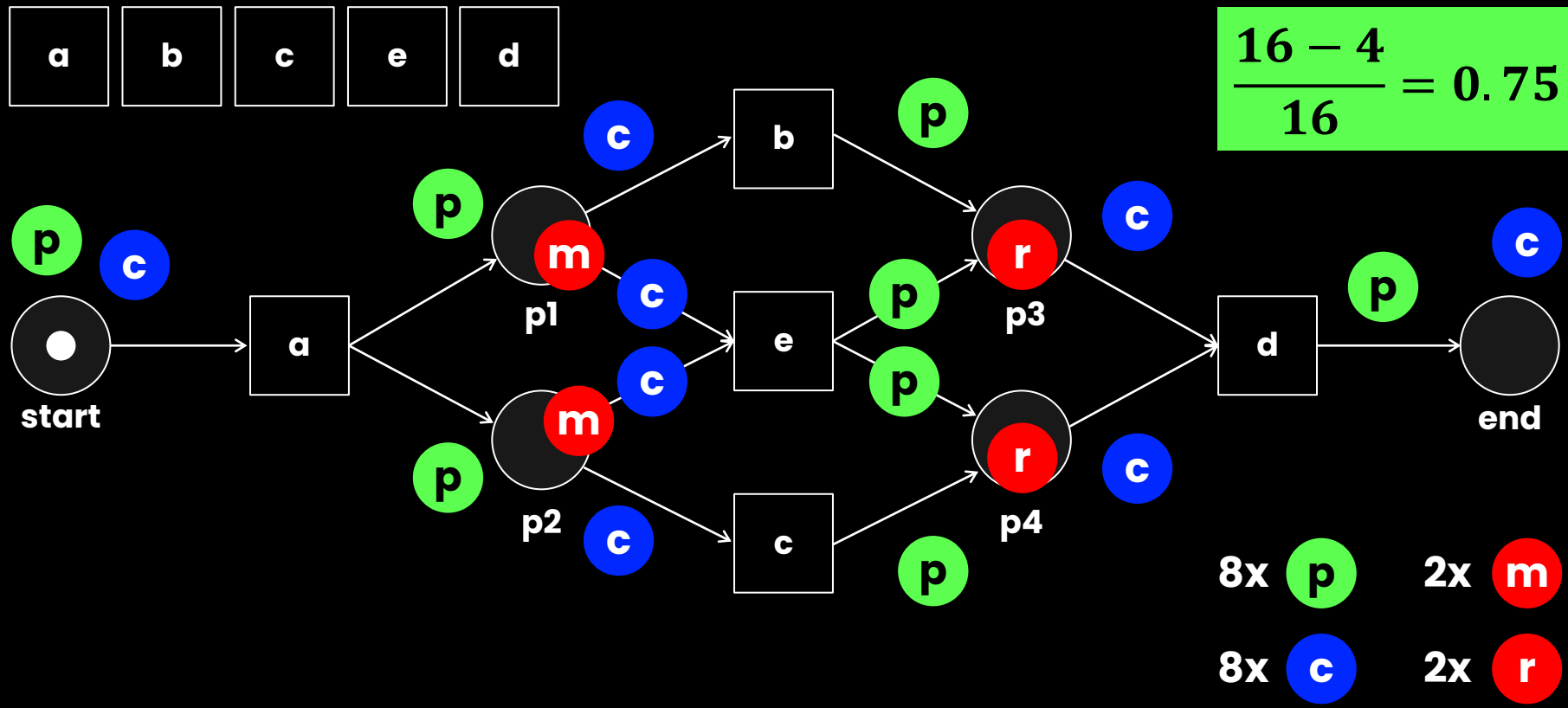
© Wil van der Aalst (use only with permission & acknowledgements)

Alternative 2: Token based replay



© Wil van der Aalst (use only with permission & acknowledgements)

Alternative 2: Token based replay

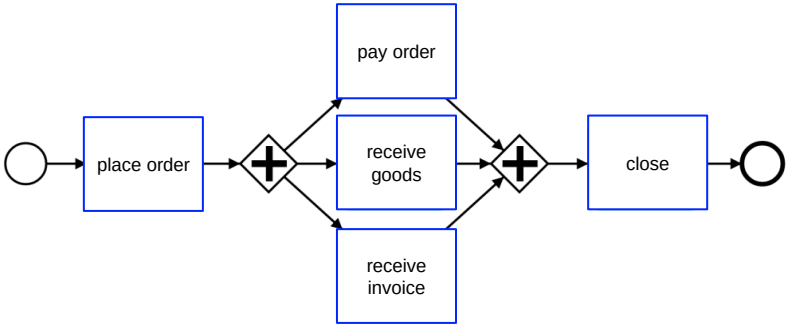


© Wil van der Aalst (use only with permission & acknowledgements)

Example: 2000 cases

853 traces 42.65%	place order	receive invoice	receive goods	pay order	close
537 traces 26.85%	place order	receive invoice	pay order	receive goods	close
537 traces 26.85%	place order	receive goods	receive invoice	pay order	close
35 traces 1.75%	place order	pay order	receive invoice	receive goods	close
27 traces 1.35%	place order	receive goods	pay order	receive invoice	close
4 traces 0,20%	place order	receive invoice	pay order	close	
4 traces 0,20%	place order	pay order	receive goods	receive invoice	close
2 traces 0.10%	receive invoice	receive goods	pay order	close	
1 traces 0.05%	receive invoice	pay order	receive goods	close	

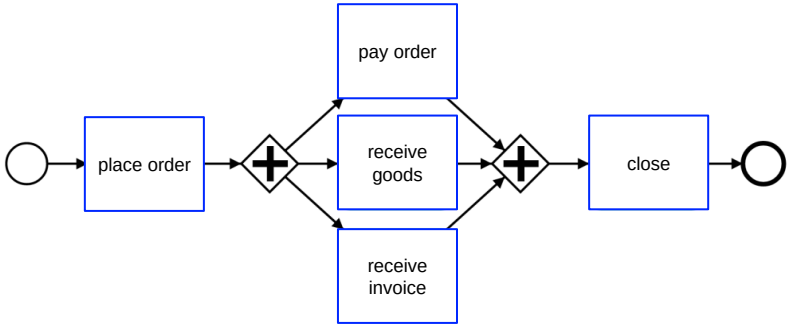
Find the deviating traces!



© Wil van der Aalst (use only with permission & acknowledgements)

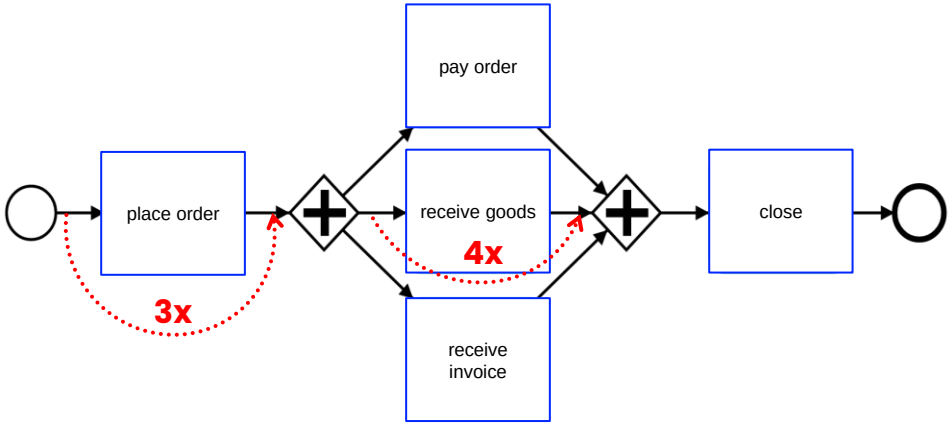
Example: 1993 cases are conforming

853 traces 42.65%	place order	receive invoice	receive goods	pay order	close
537 traces 26.85%	place order	receive invoice	pay order	receive goods	close
537 traces 26.85%	place order	receive goods	receive invoice	pay order	close
35 traces 1.75%	place order	pay order	receive invoice	receive goods	close
27 traces 1.35%	place order	receive goods	pay order	receive invoice	close
4 traces 0,20%	place order	pay order	receive goods	receive invoice	close



© Wil van der Aalst (use only with permission & acknowledgements)

Example: 7 cases are not conforming



4x

missing re-ceived goods

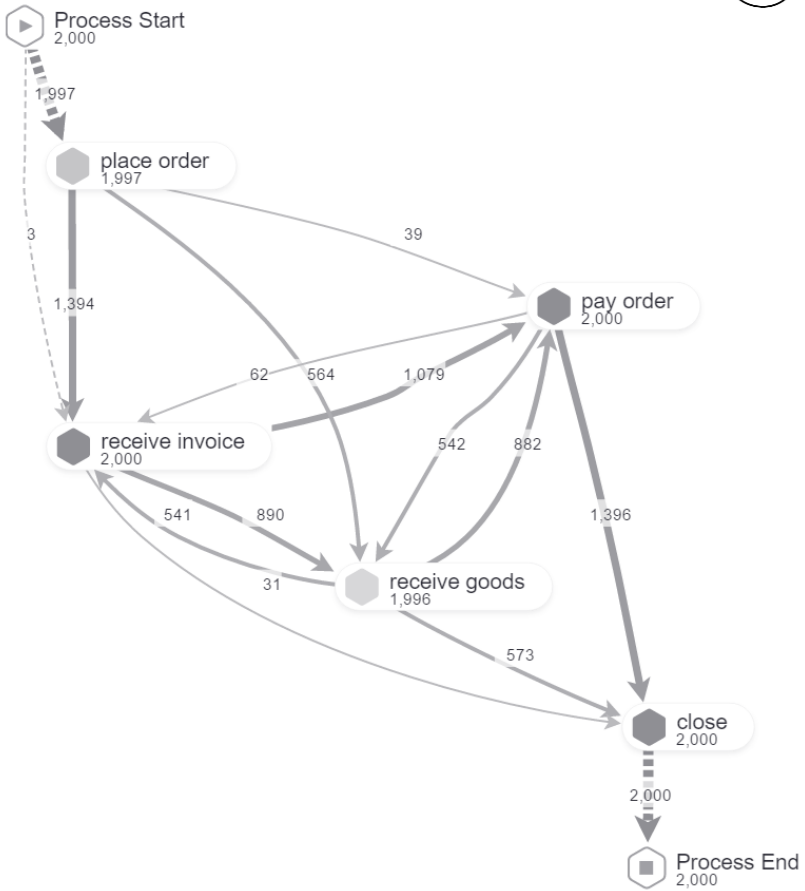
3x

missing place order

© Wil van der Aalst (use only with permission & acknowledgements)

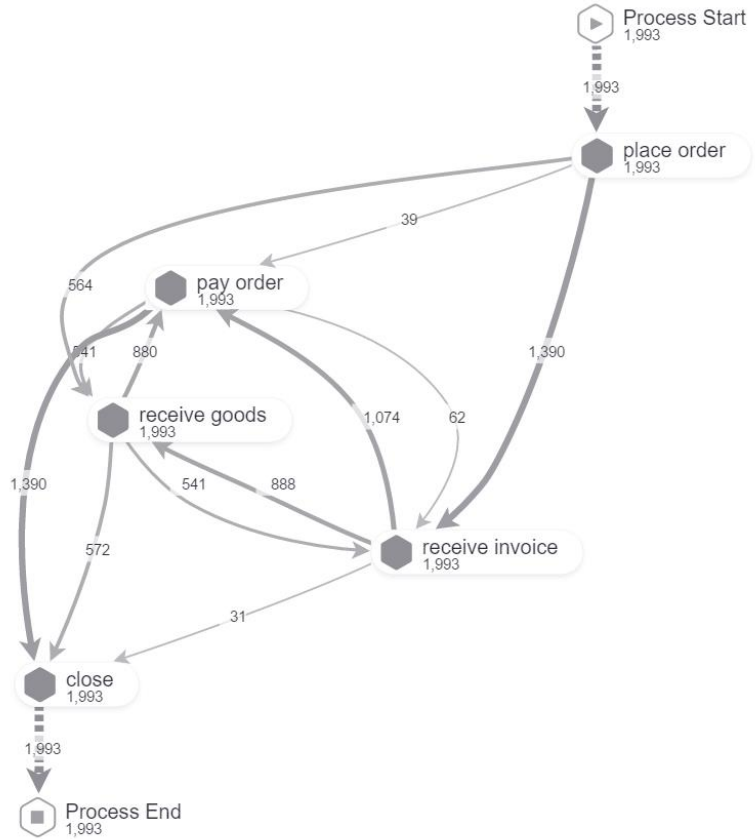
Not so easy to see the deviations in a DFG

853 traces 42.65%	place order	receive invoice	receive goods	pay order	close
537 traces 26.85%	place order	receive invoice	pay order	receive goods	close
537 traces 26.85%	place order	receive goods	receive invoice	pay order	close
35 traces 1.75%	place order	pay order	receive invoice	receive goods	close
27 traces 1.35%	place order	receive goods	pay order	receive invoice	close
4 traces 0,20%	place order	receive invoice	pay order	close	
4 traces 0,20%	place order	pay order	receive goods	receive invoice	close
2 traces 0.10%	receive invoice	receive goods	pay order	close	
1 traces 0.05%	receive invoice	pay order	receive goods	close	



© Wil van der Aalst (use only with permission & acknowledgements)

DFG of conforming cases

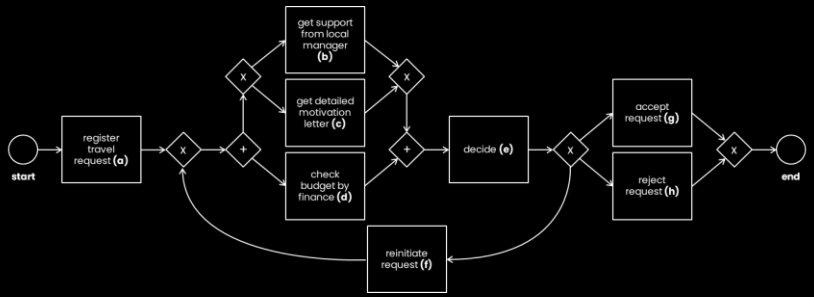


© Wil van der Aalst (use only with permission & acknowledgements)

DFG of deviating cases

© Wil van der Aalst (use only with permission & acknowledgements)

Conformance checking: Summary



conforming

a b d e g

deviating

a b c e h

Relevance

- To answer compliance questions
- Understand where model and reality disagree
- Auditing, fraud detection, process improvement

We have seen three techniques

- Alignments
- Footprint comparison
- Token-based replay

© Wil van der Aalst (use only with permission & acknowledgements)

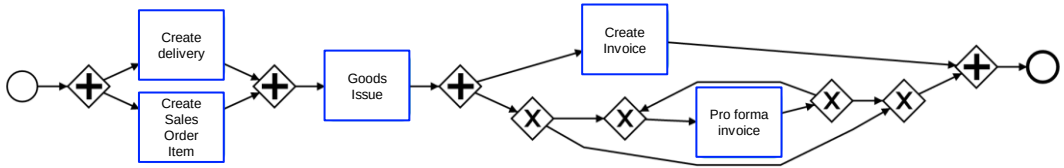
Token based replay in Celonis

- Celonis is using a variant of token-based replay
- The BPMN model is translated to a Petri net
- Diagnostics are presented as sentences

Conforming cases
60%
Cases conforming

Conforming cases
1.97k vs. **1.34k**
Cases conforming

Throughput time
32.4 vs. **7.6** Days
Violations increased throughput time by 24.8 Days



1% of cases

Delivery Block changed is an undesired activity

Add to whitelist View cases in...

Effect on throughput time	Effect on steps per case
64 Days longer	+ 11.5 Steps per case

1% of cases

Create Sales Order Item is followed by *Create Invoice*

Add to whitelist View cases in...

Effect on throughput time	Effect on steps per case
4 Days longer	- 1.0 Steps per case

1% of cases

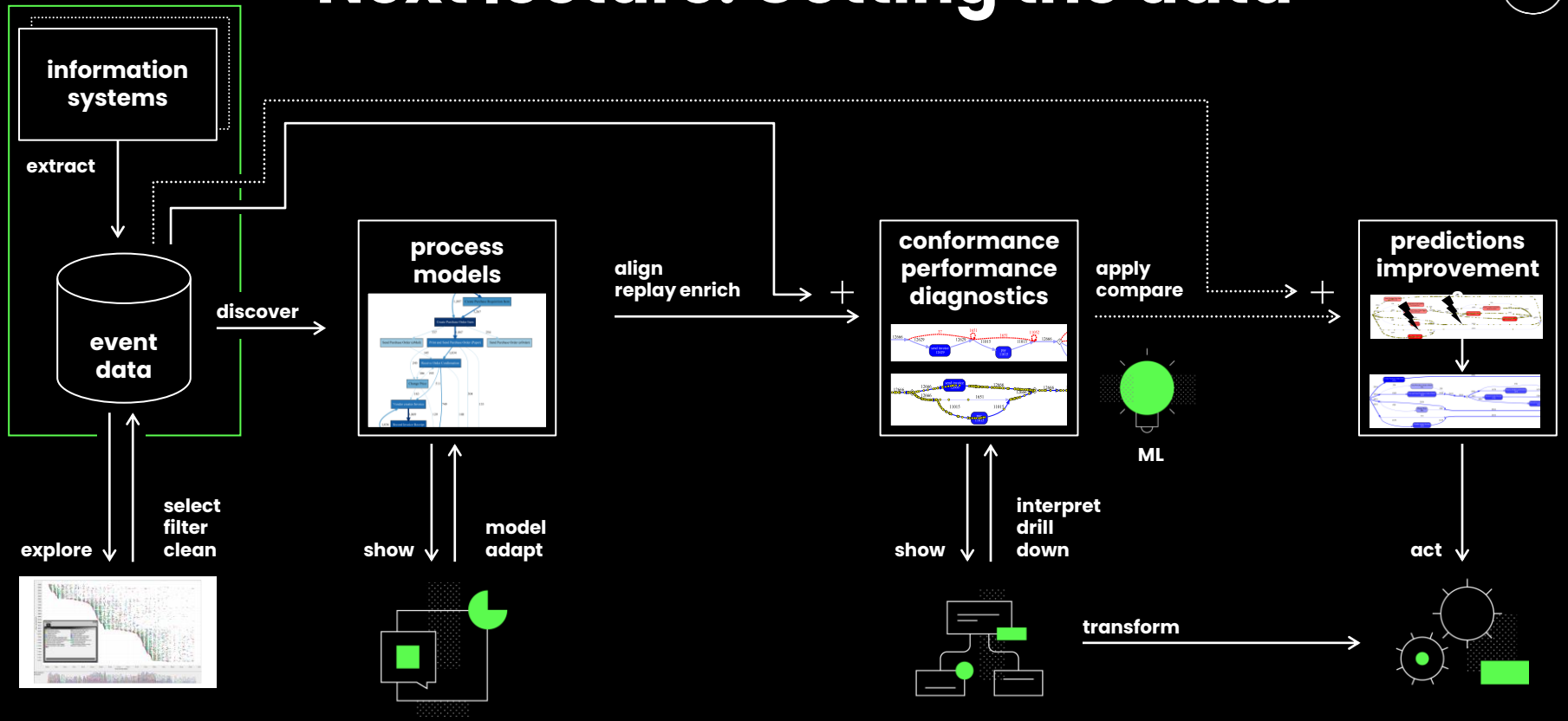
Goods Issue is followed by *Create Delivery*

Add to whitelist View cases in...

Effect on throughput time	Effect on steps per case
49 Days longer	+ 8.4 Steps per case

© Wil van der Aalst (use only with permission & acknowledgements)

Next lecture: Getting the data



© Wil van der Aalst (use only with permission & acknowledgements)