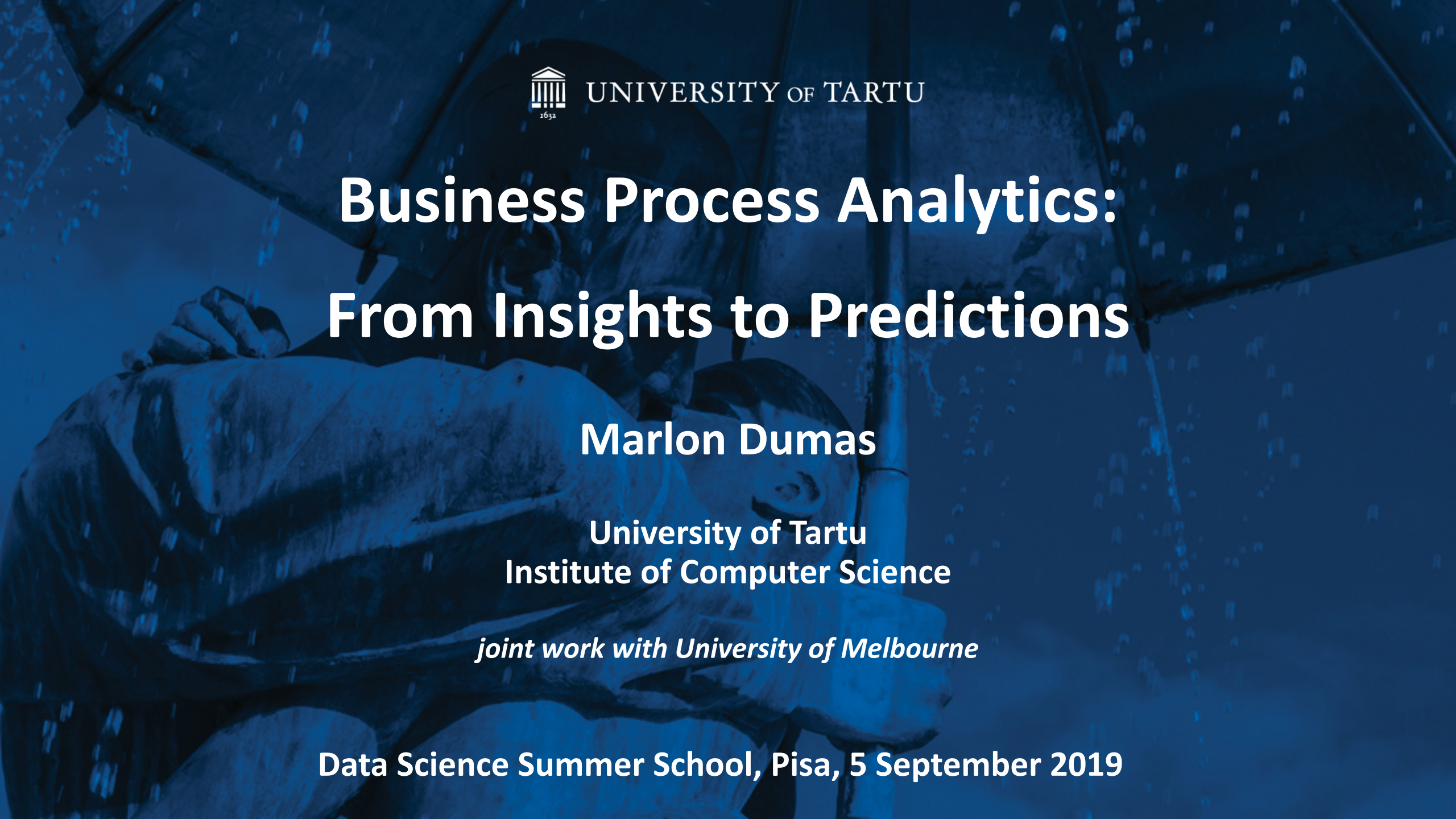


UNIVERSITY OF TARTU

Business Process Analytics: From Insights to Predictions

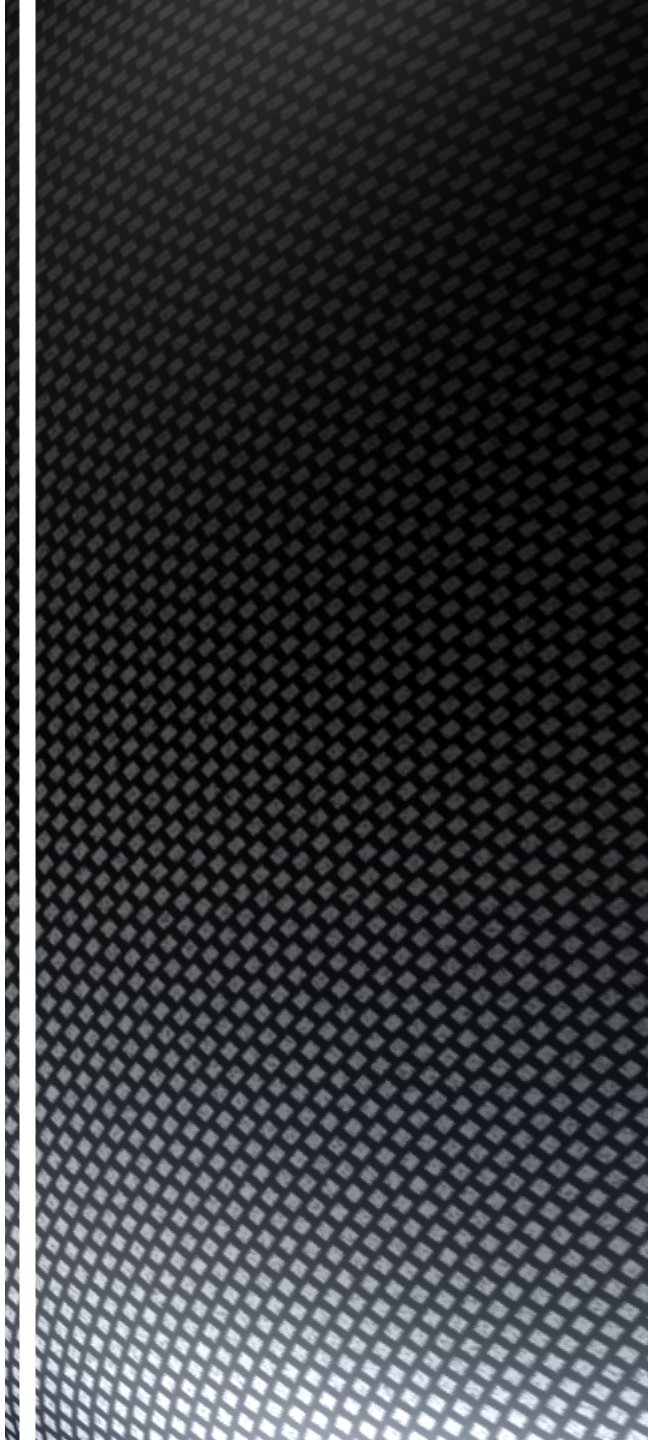
Marlon Dumas

University of Tartu
Institute of Computer Science

joint work with University of Melbourne

Data Science Summer School, Pisa, 5 September 2019

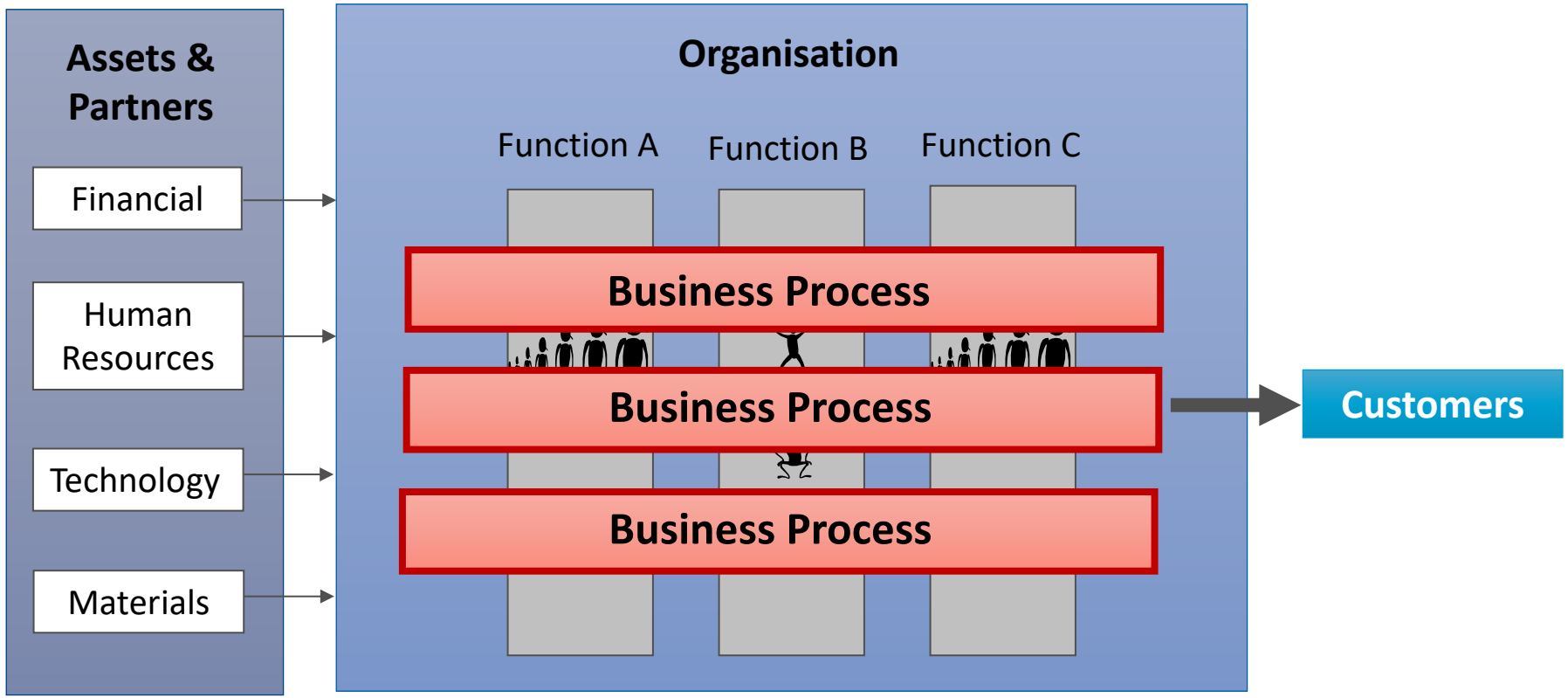
What is a Business Process?

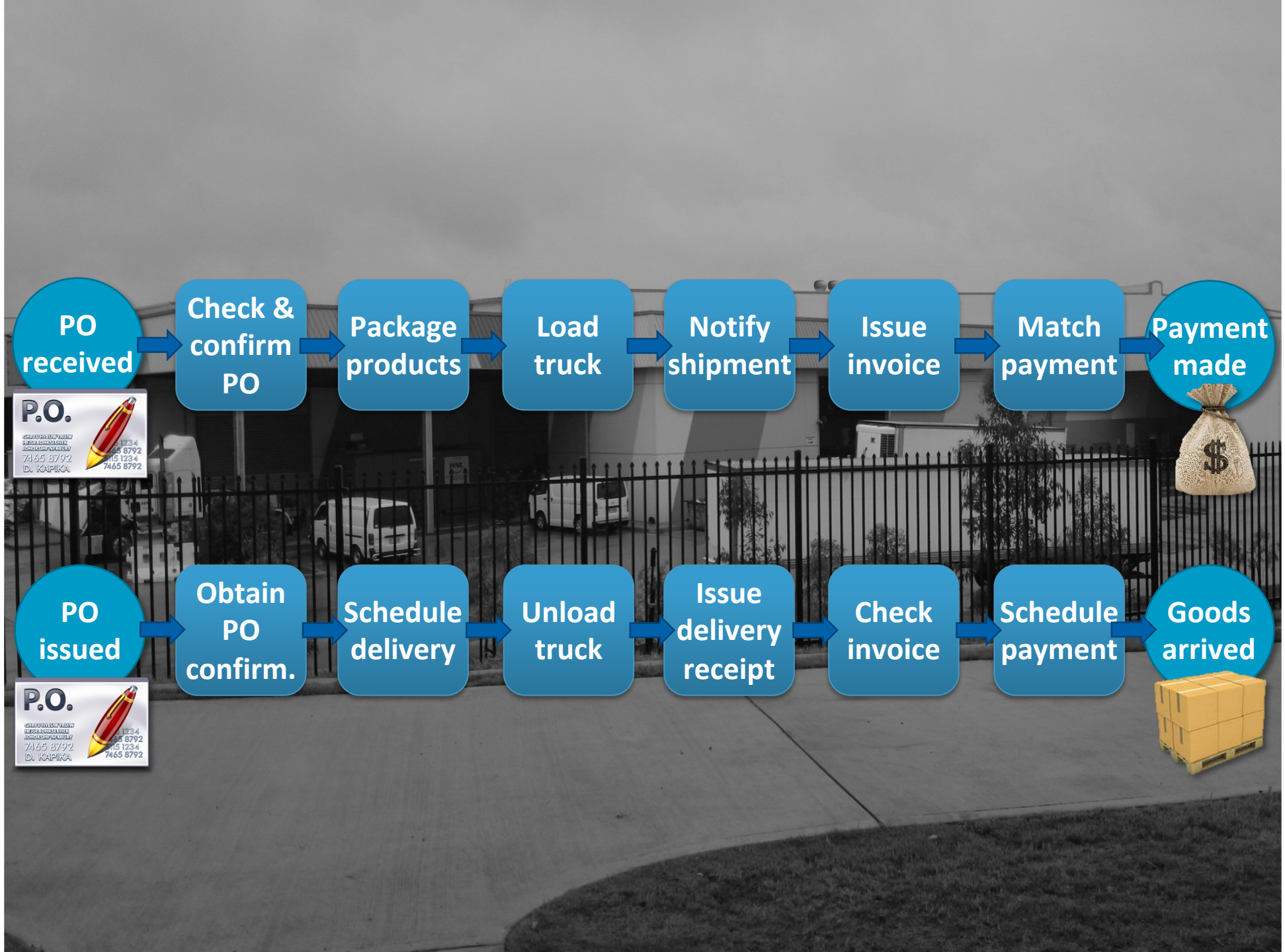


3


Issue delivery receipt
Load truck
Package products
Issue invoice
Prepare shipment
Schedule payment
Schedule delivery
Check & confirm PO
Unload truck
Notify shipment
Obtain PO confirm.
Check Invoice
Request PO change
Match incoming payment

Business processes




PO received
Check & confirm PO
Package products
Load truck
Notify shipment
Issue invoice
Match payment
Payment made
P.O.
PO issued
Obtain PO confirm.
Schedule delivery
Unload truck
Issue delivery receipt
Check invoice
Schedule payment
Goods arrived
P.O.

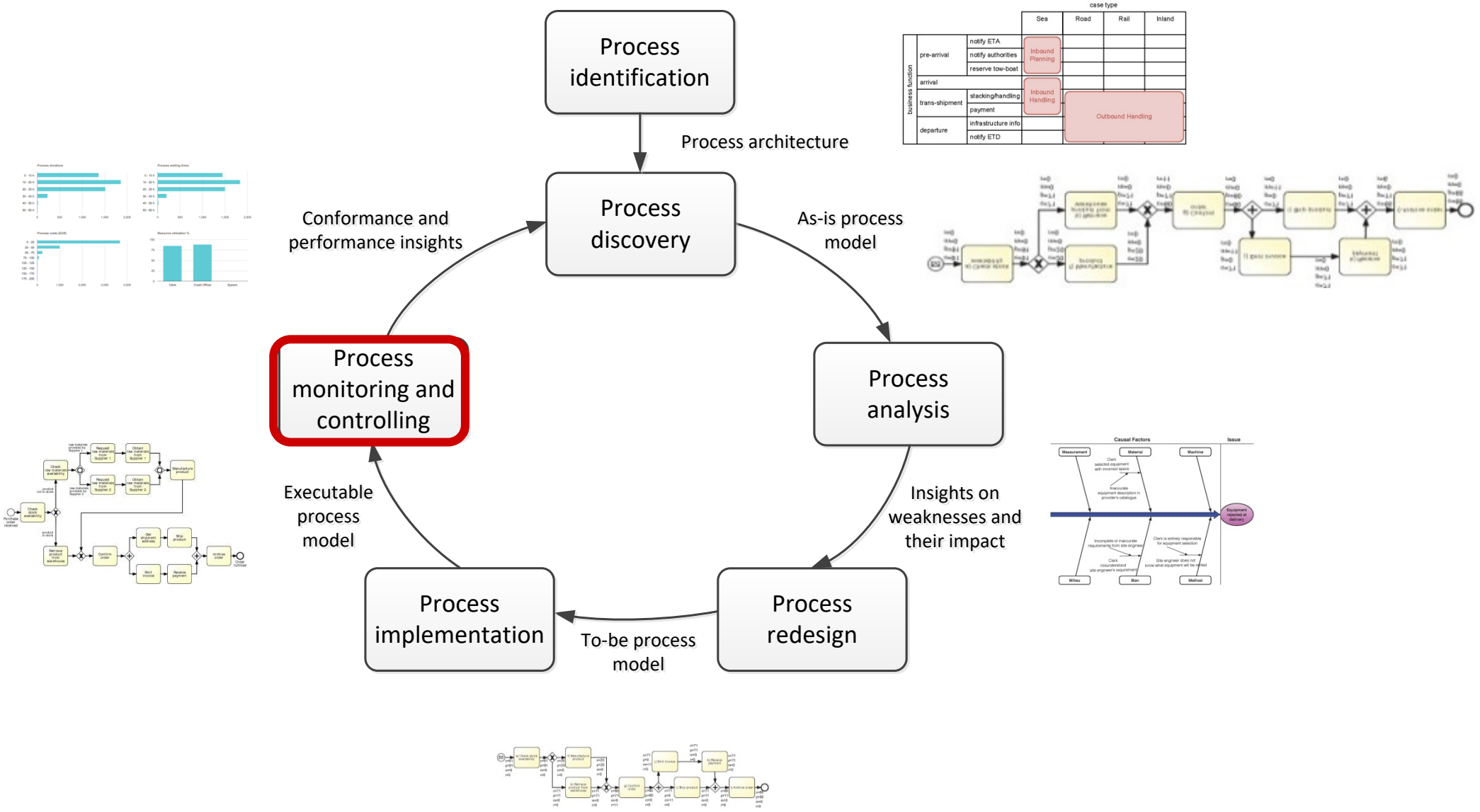
A business process is…

a chain of events, activities and decisions
…involving a number of actors and objects,
….triggered by a need
and leading to an outcome that is of value to a customer.

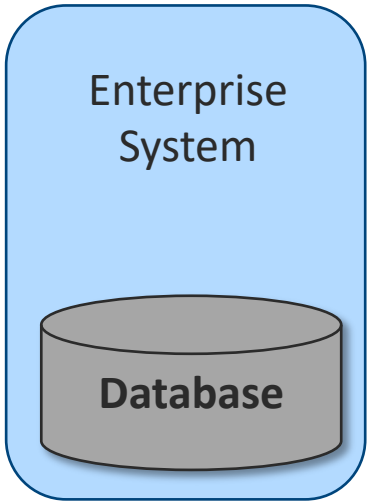
Examples:

- Order-to-Cash
- Procure-to-Pay (aka Purchase-to-Pay)
- Application-to-Approval
- Fault-to-Resolution

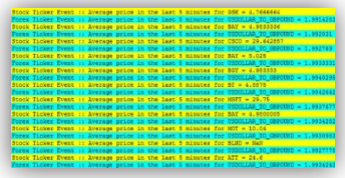
The Business Process Management (BPM) lifecycle



Business Process Monitoring

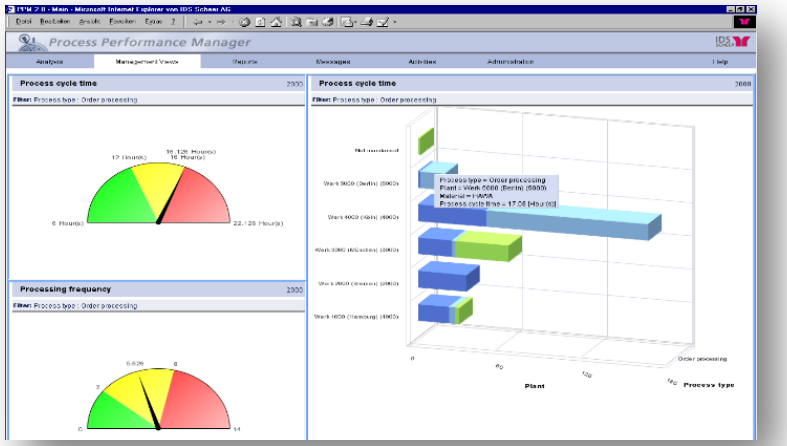


Event stream

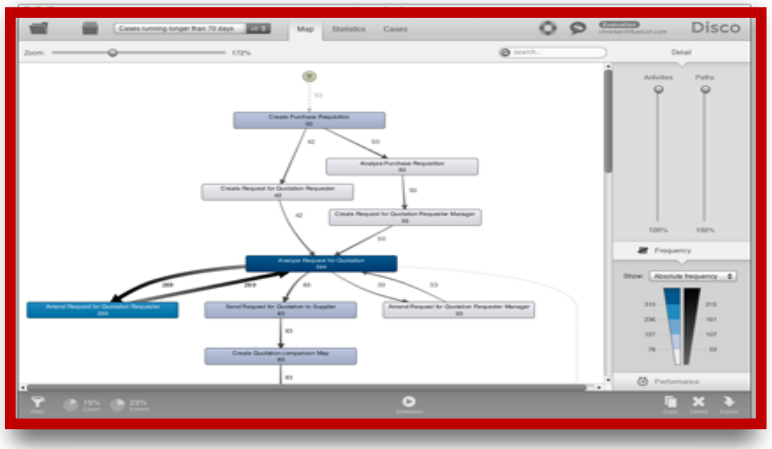


Event log

Performance Dashboards



Process Mining



Structure of a Business Process Event Log

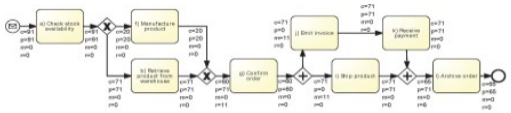
Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
			a	Mortgage	30 000	-
			a	Mortgage	30 000	-
			a	Mortgage	30 000	-
...	...	...	...	...	...	...

Concrete formats:

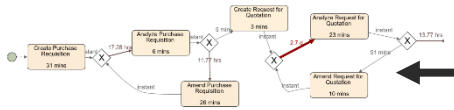
- Comma-Separated Values (CSV)
- IEEE XES (XML format)

Tactical Process Mining

Automated Process Discovery



discovered process model

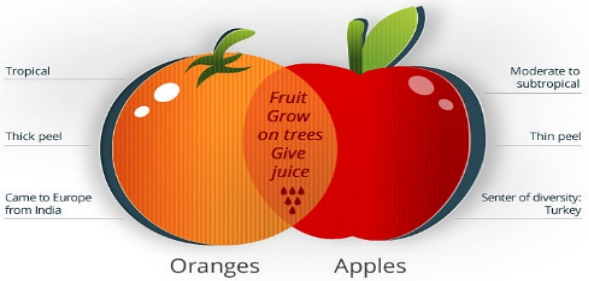


Enhanced process model

Performance Mining

event log

Variants Analysis

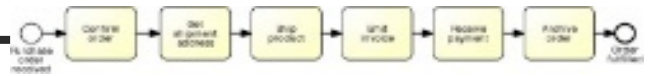


Difference diagnostics

event log'

Conformance Checking

✓ / ✗

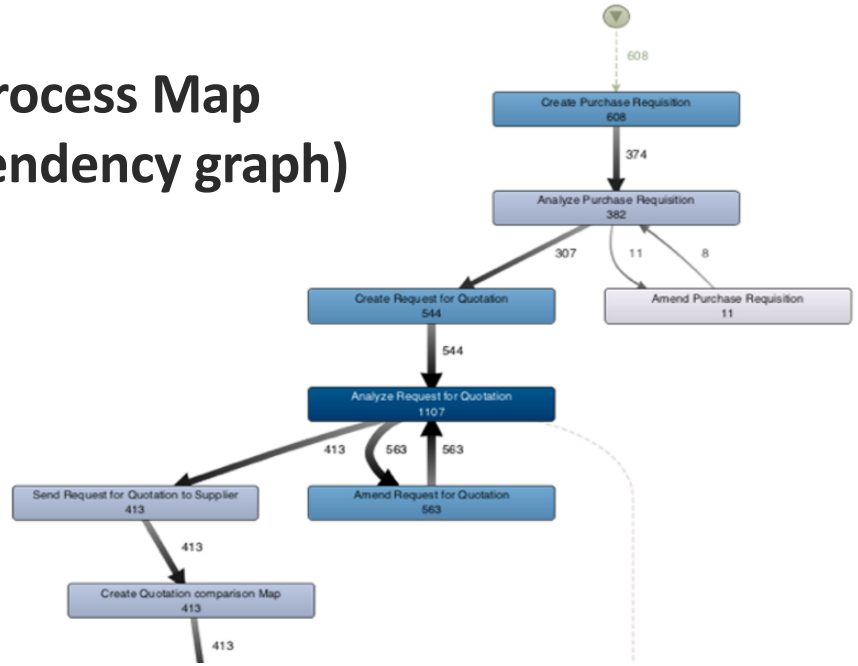


input process model

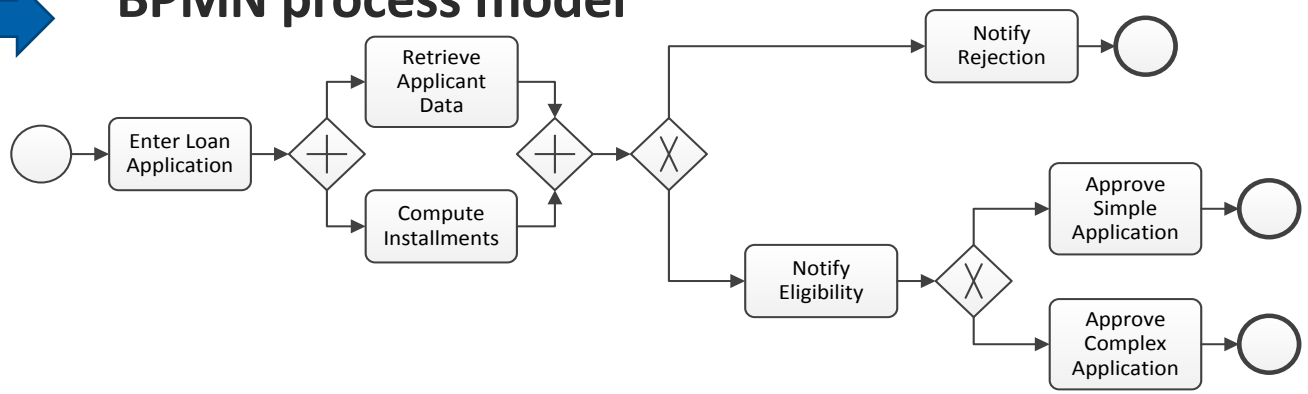
Automated Process Discovery

CID	Task	Time Stamp	…
13219	Enter Loan Application	2007-11-09 T 11:20:10	-
13219	Retrieve Applicant Data	2007-11-09 T 11:22:15	-
13220	Enter Loan Application	2007-11-09 T 11:22:40	-
13219	Compute Installments	2007-11-09 T 11:22:45	-
13219	Notify Eligibility	2007-11-09 T 11:23:00	-
13219	Approve Simple Application	2007-11-09 T 11:24:30	-
13220	Compute Installements	2007-11-09 T 11:24:35	-
…	…	…	…

Process Map (dependency graph)



BPMN process model



Discovering BPMN Process Models

Alpha miner (α-miner)

- *Simple,* limited, not robust

Heuristics miner (and derivatives, including Fodina)

- Robust to noise, fast, but can produce incorrect models

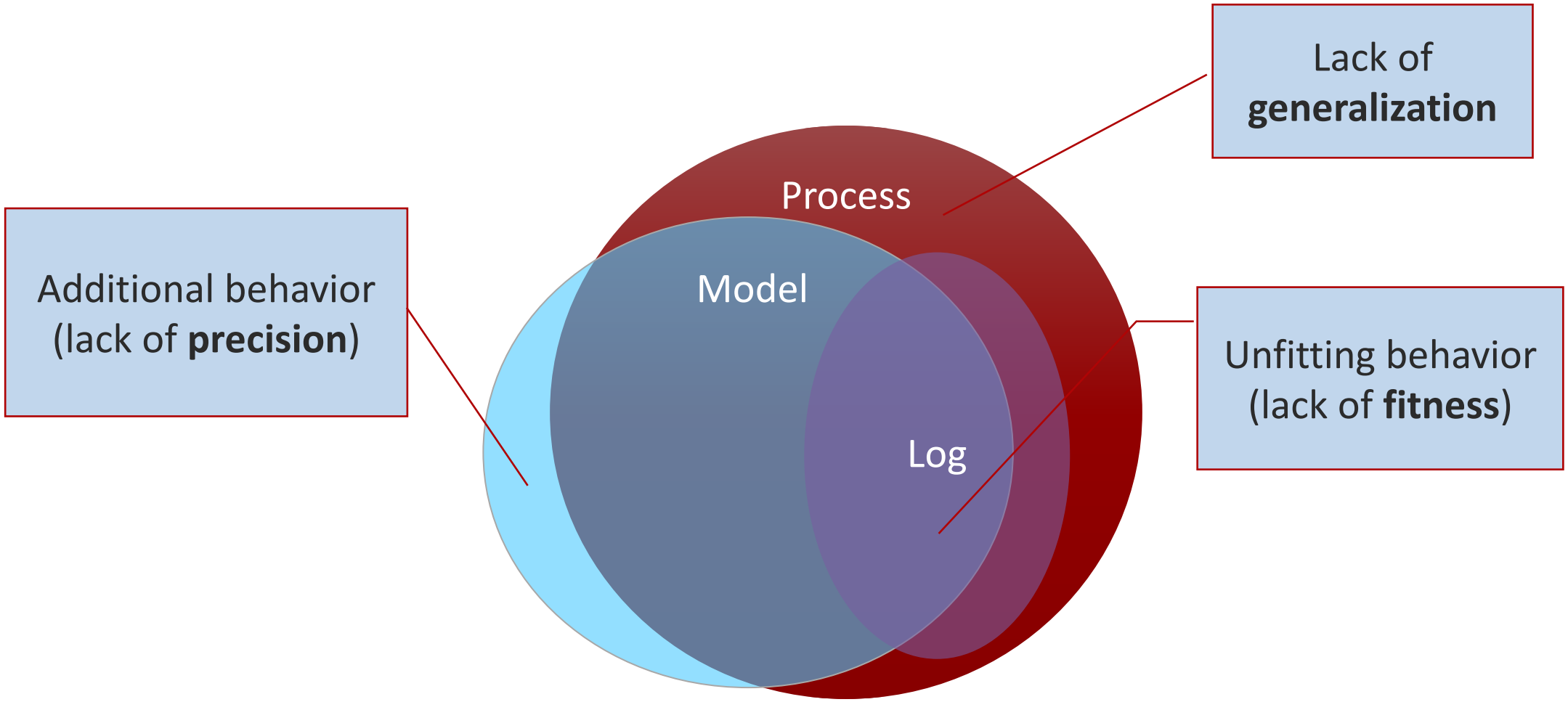
Inductive miner

- Ensures that models are block-structured & correct

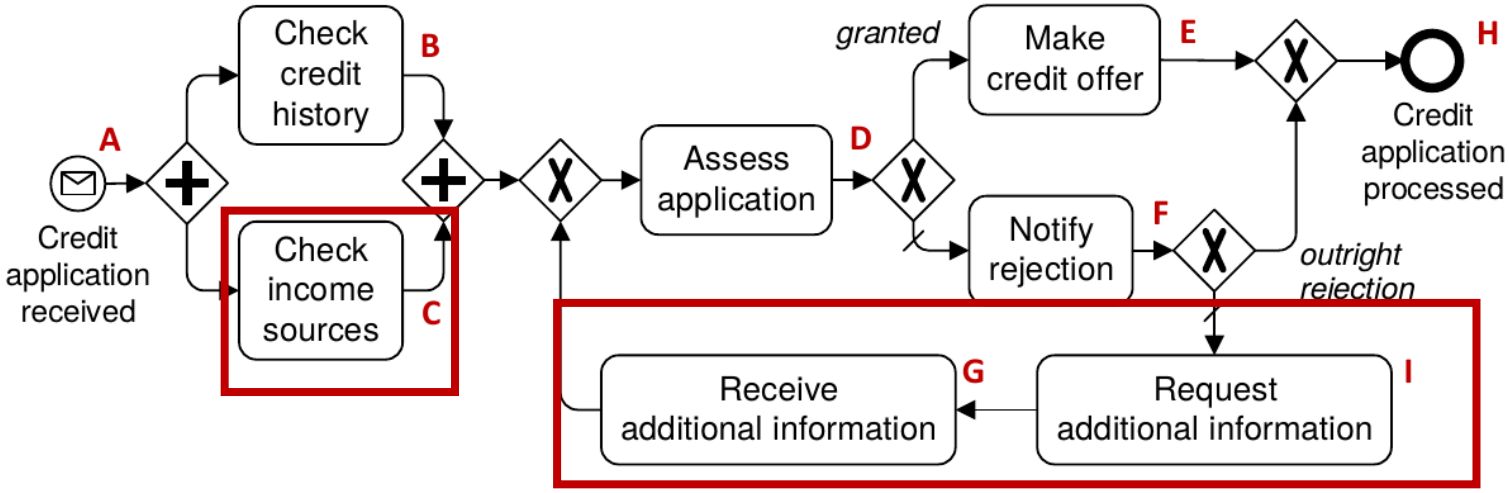
Split miner

- Produces deadlock-free but not necessarily structured models

Accuracy of Automated Process Discovery



Conformance Checking in Apromore



Event log:

ABCDEH

ACBDEH

ABCDFH

ACBDFH

ABDEH

ABDFH

Lack of fitness

Lack of precision

García-Bañuelos et al. “Complete and Interpretable Conformance Checking of Business Processes” *IEEE Transactions on Software Engineering* 44(3): 262-290, 2018

Automated Process Discovery Benchmark

- 24 real-life event logs (most from IEEE Task force on Process Mining)
- Quality criteria:
  - Accuracy measures: Fitness, precision, F-Score, generalization
  - Model complexity measures: size, structural complexity, structuredness
  - Model soundness
  - Execution time
- Main conclusions:
  - Inductive Miner, Evolutionary Tree Miner, Split Miner have highest F-scores
    - Closely followed by Fodina
  - Inductive Miner achieves highest fitness generally, but lower precision (than Split Miner)
  - Evolutionary tree miner produces simpler models, but high execution times

Augusto et al. “Automated Discovery of Process Models from Event Logs: Review and Benchmark”. *IEEE Transactions on Knowledge and Data Engineering* (2018), DOI: 10.1109/TKDE.2018.2841877

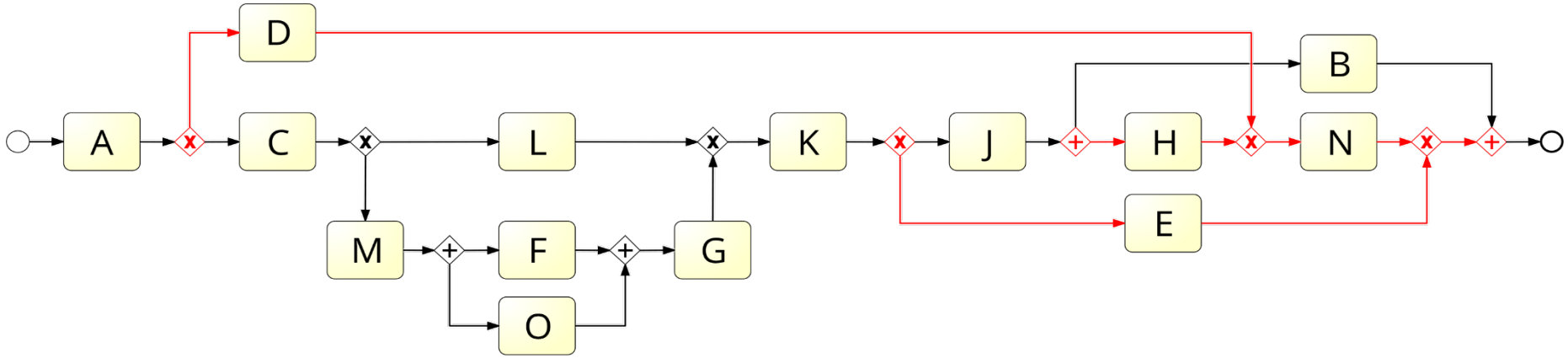
Automated Process Discovery Methods

Heuristics Miner

good F-score

complex models

semantic errors



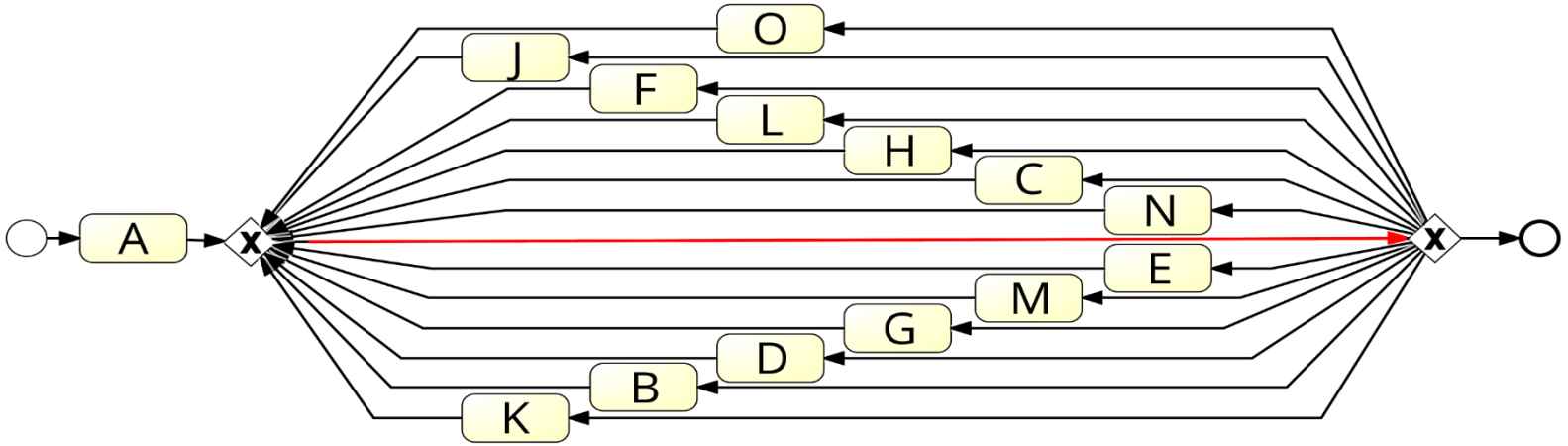
Inductive Miner

high fitness

no semantic errors

simpler models

low precision



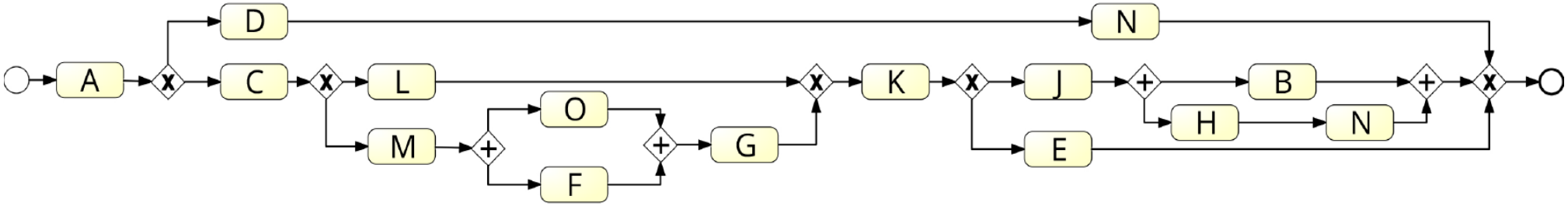
Split Miner

high fitness

no semantic errors

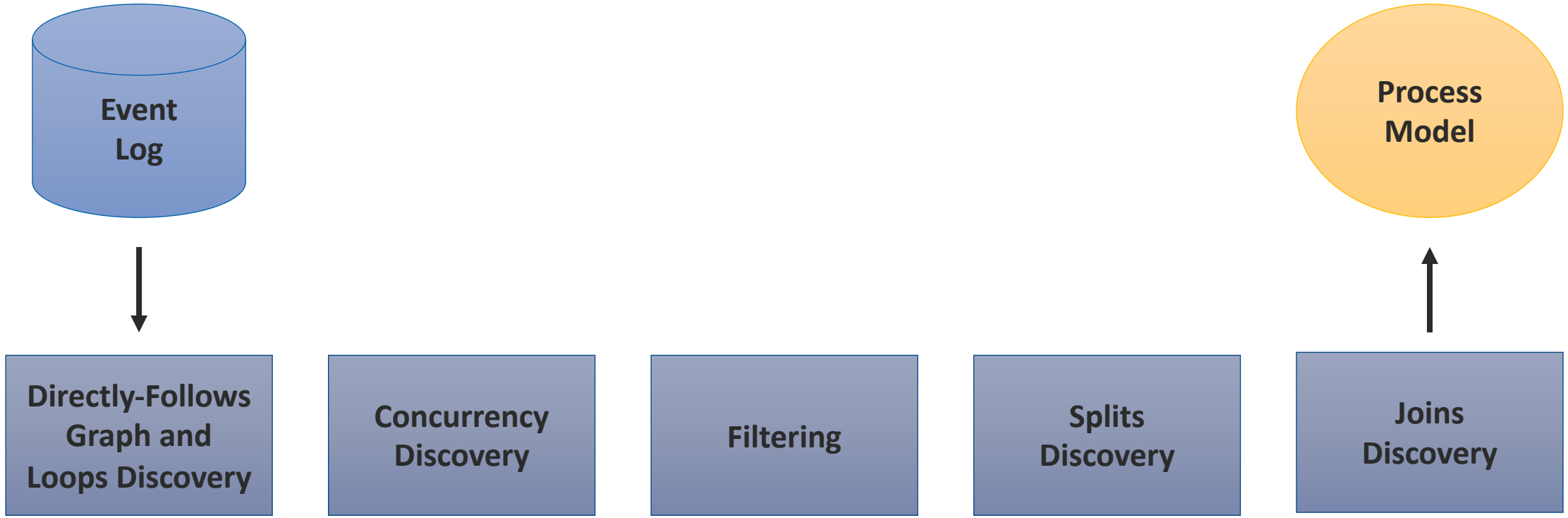
simpler models

Moderate precision



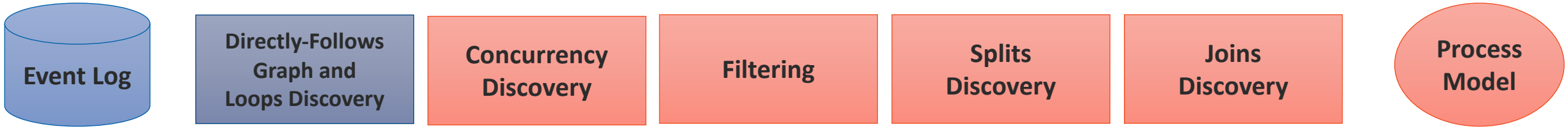
A. Augusto et al. Split Miner: Discovering Accurate and Simple Business Process Models from Event Logs. In *ICDM*'2017.

Split Miner Algorithm

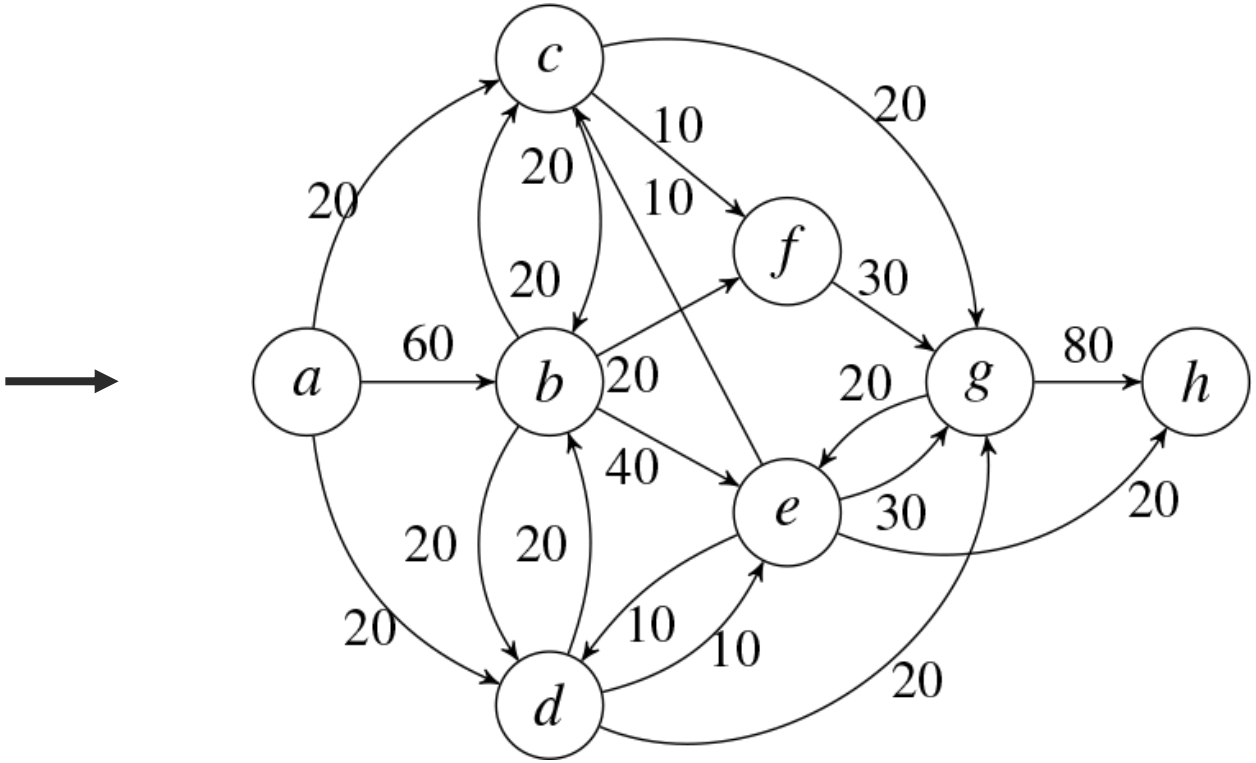


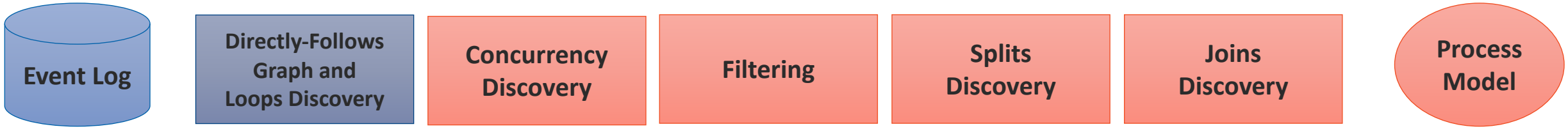
- Adriano Augusto, Raffaele Conforti, Marlon Dumas, Marcello La Rosa, Artem Polyvyanyy: Split miner: automated discovery of accurate and simple business process models from event logs. Knowl. Inf. Syst. 59(2): 251-284 (2019)

Split Miner

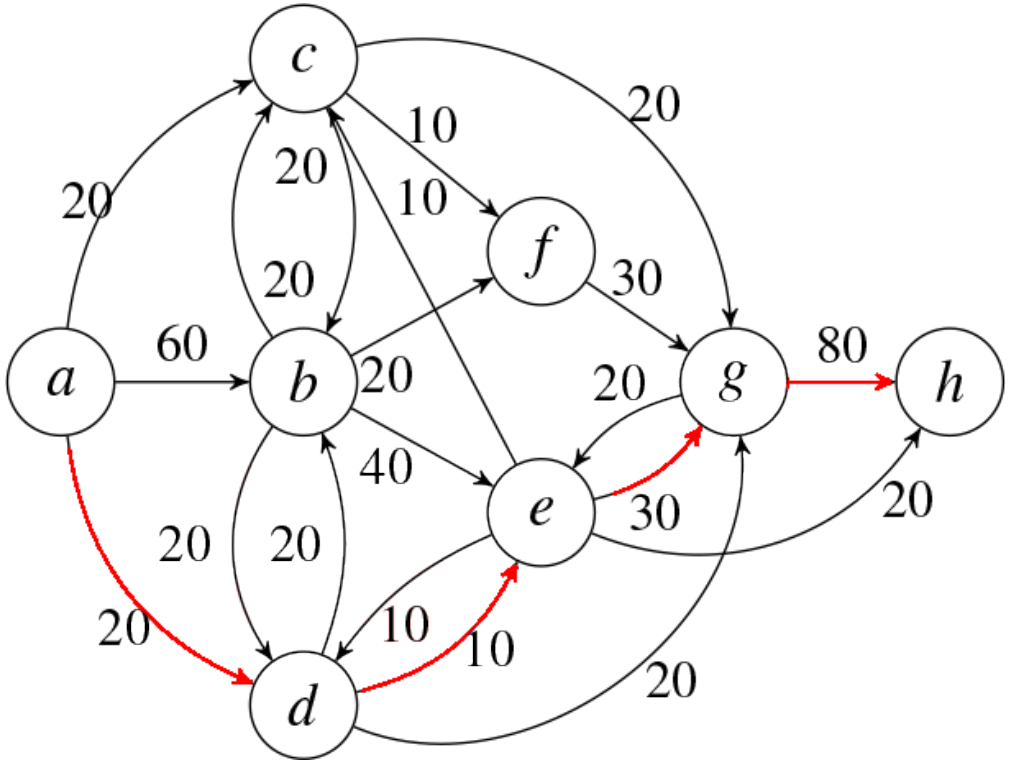


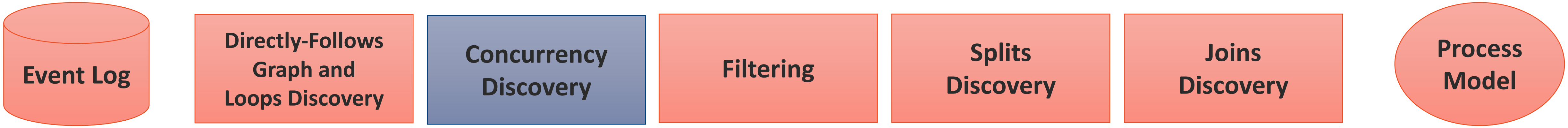
Trace	#obs
a » b » c » g » e » h	10
a » b » c » f » g » h	10
a » b » d » g » e » h	10
a » b » d » e » g » h	10
a » b » e » c » g » h	10
a » b » e » d » g » h	10
a » c » b » e » g » h	10
a » c » b » f » g » h	10
a » d » b » e » g » h	10
a » d » b » f » g » h	10

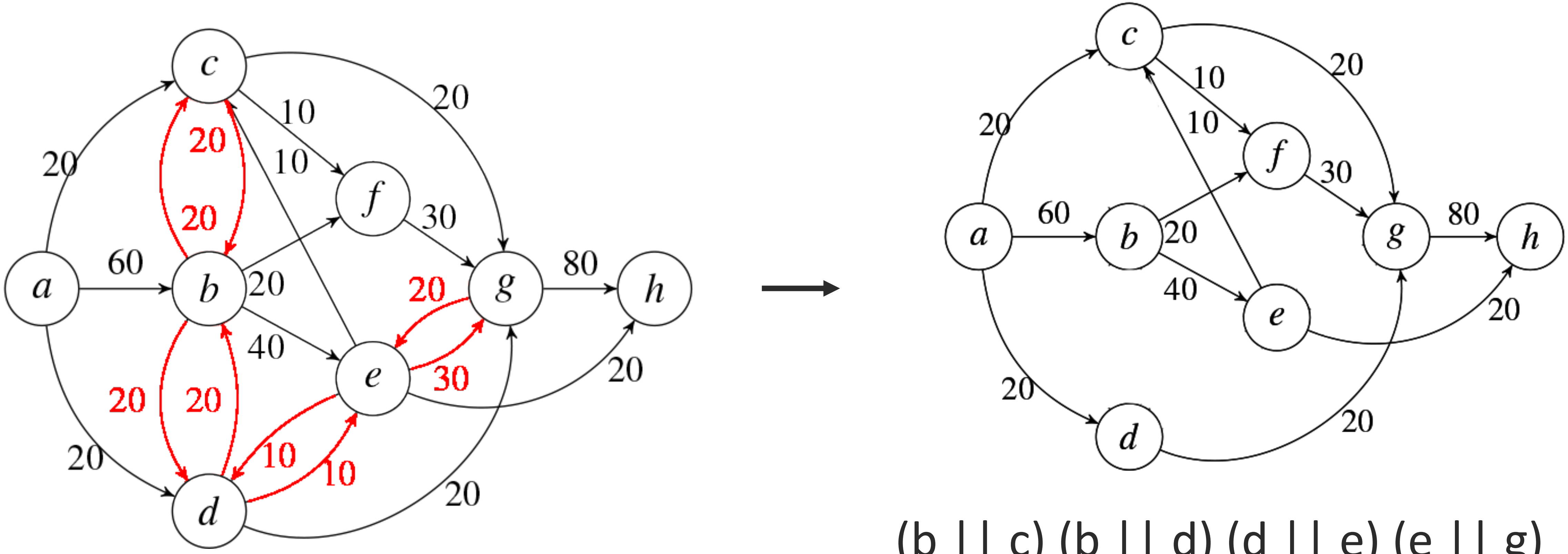




Trace	#obs
a » b » c » g » e » h	10
a » b » c » f » g » h	10
a » b » d » g » e » h	10
a » b » d » e » g » h	10
a » b » e » c » g » h	10
a » b » e » d » g » h	10
a » c » b » e » g » h	10
a » c » b » f » g » h	10
a » d » b » e » g » h	10
a » d » b » f » g » h	10

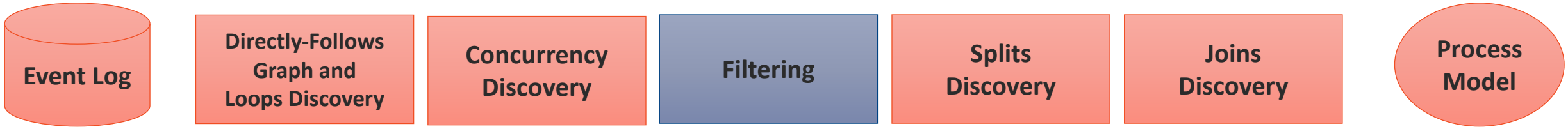


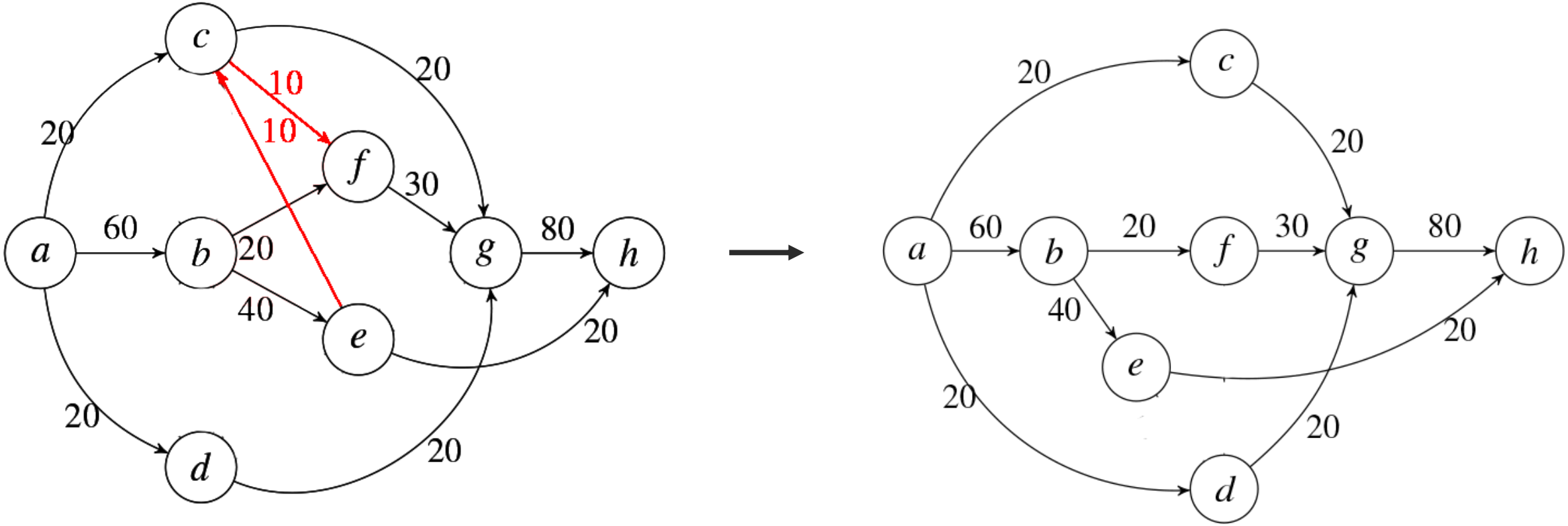

Event Log
Directly-Follows Graph and Loops Discovery
Concurrency Discovery
Filtering
Splits Discovery
Joins Discovery
Process Model


a
b
c
d
e
f
g
h
60
80
40
30
20
10

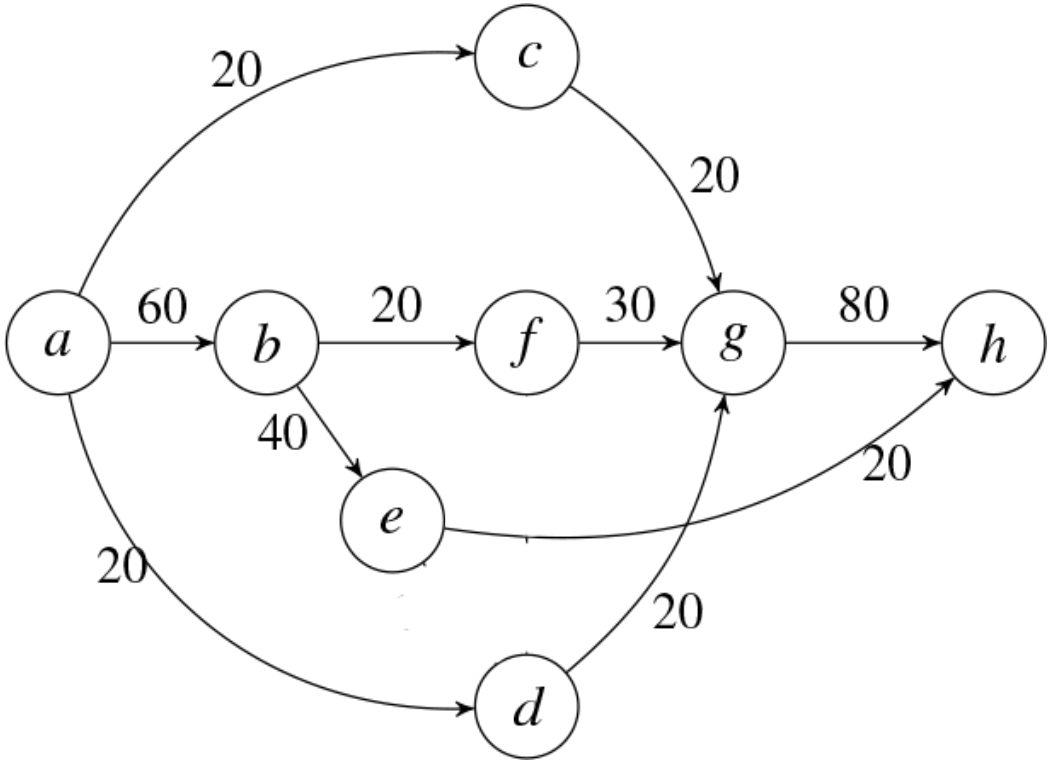
(b || c) (b || d) (d || e) (e || g)

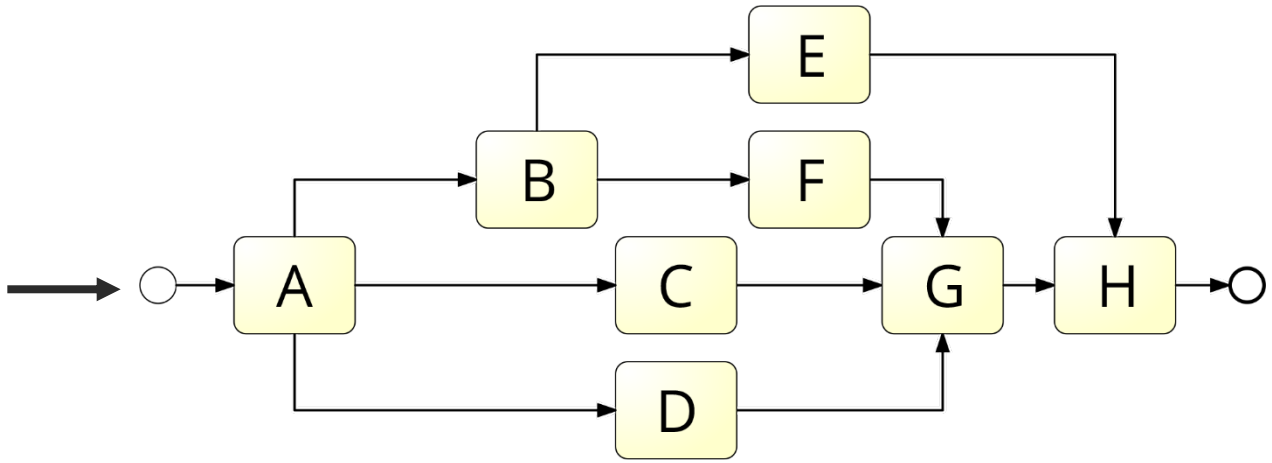
Split Miner


Event Log
Directly-Follows Graph and Loops Discovery
Concurrency Discovery
Filtering
Splits Discovery
Joins Discovery
Process Model

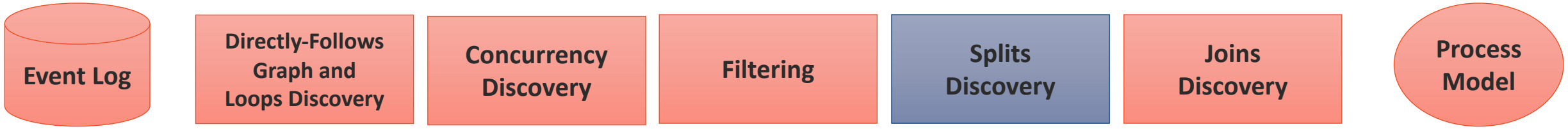

a
b
c
d
e
f
g
h
20
60
10
10
20
30
80
20
40
20
20
20
20
60
20
20
30
80
40
20
20
20

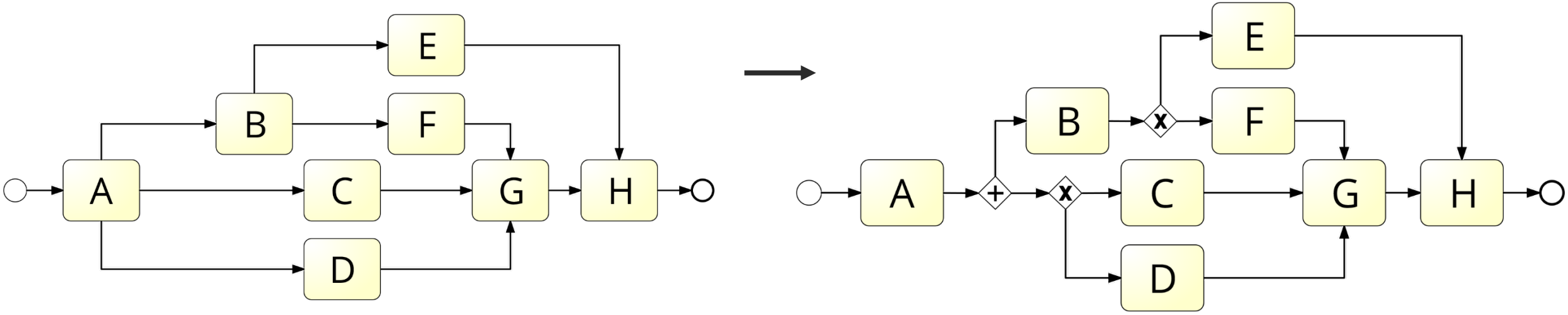
From Maps to BPMN





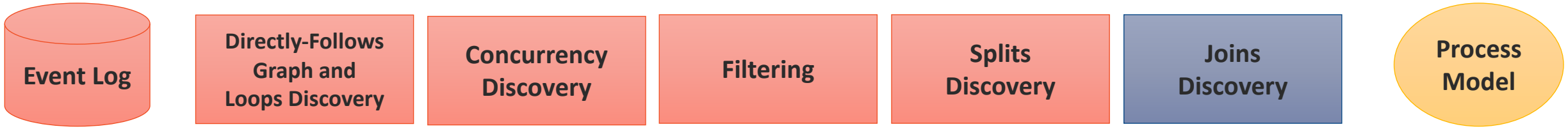
Split Miner

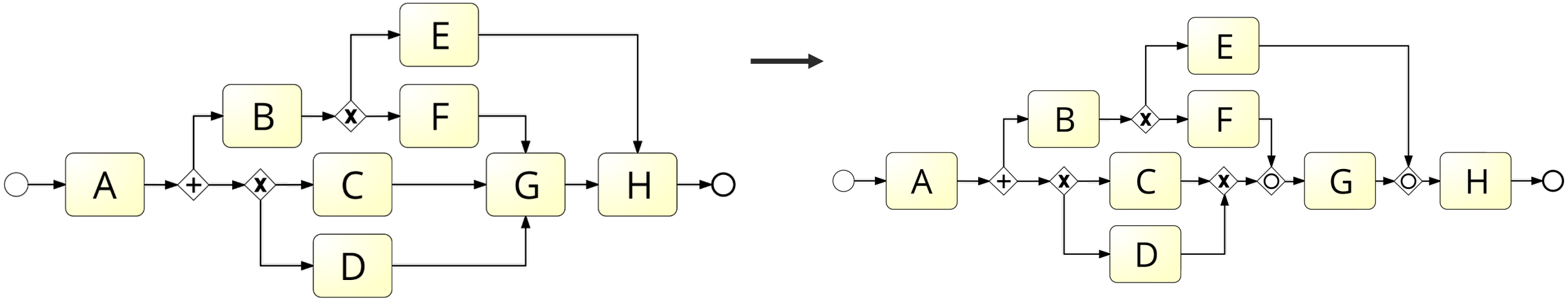




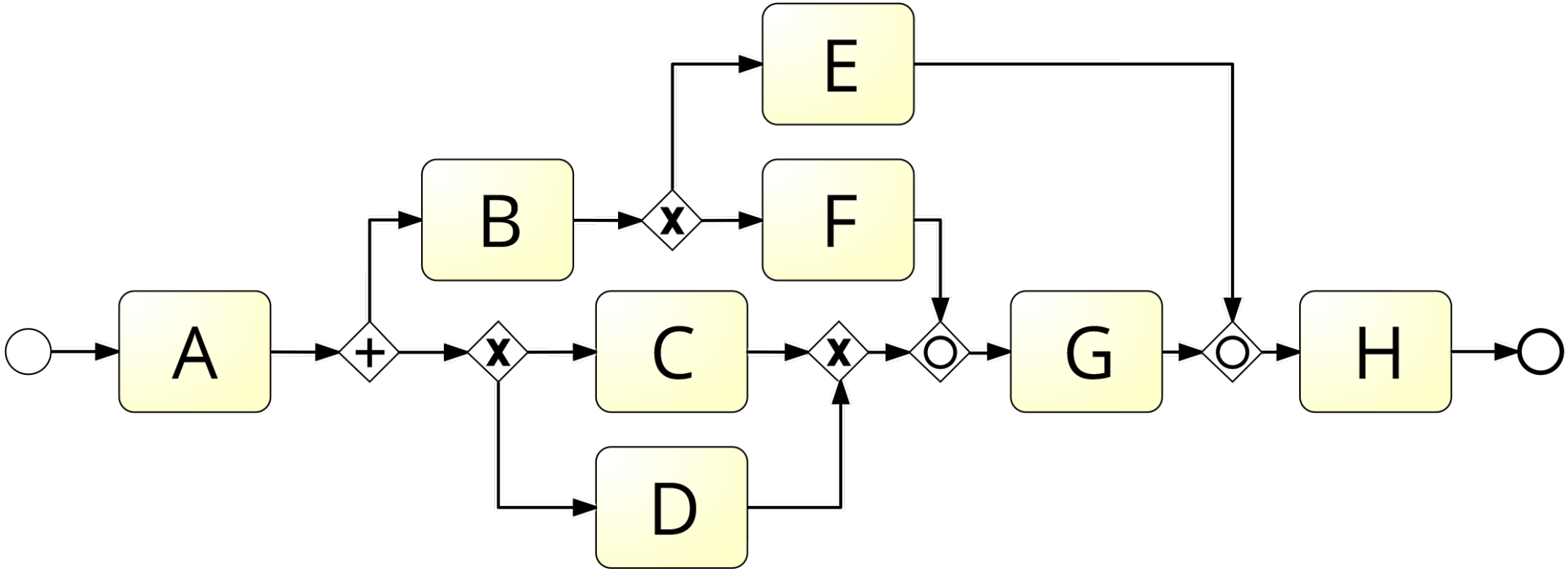
(b || c) (b || d) (d || e) (e || g)

Split Miner


Event Log
Directly-Follows Graph and Loops Discovery
Concurrency Discovery
Filtering
Splits Discovery
Joins Discovery
Process Model


A
B
C
D
E
F
G
H
+
x
x
A
B
C
D
E
F
G
H
+
x
x
o
o

Done!

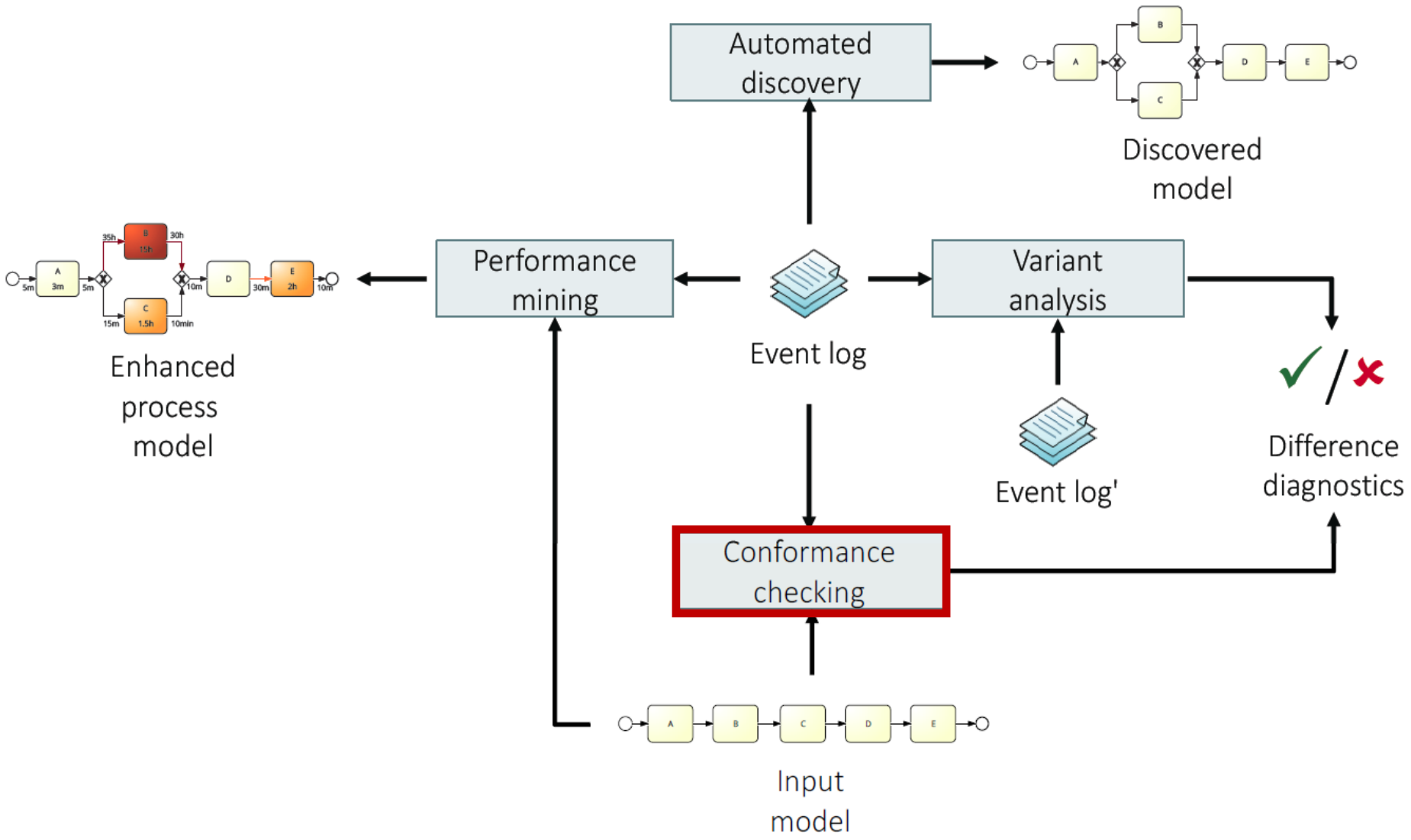


UNIVERSITY OF TARTU

Demo Time!

http://apromore.org

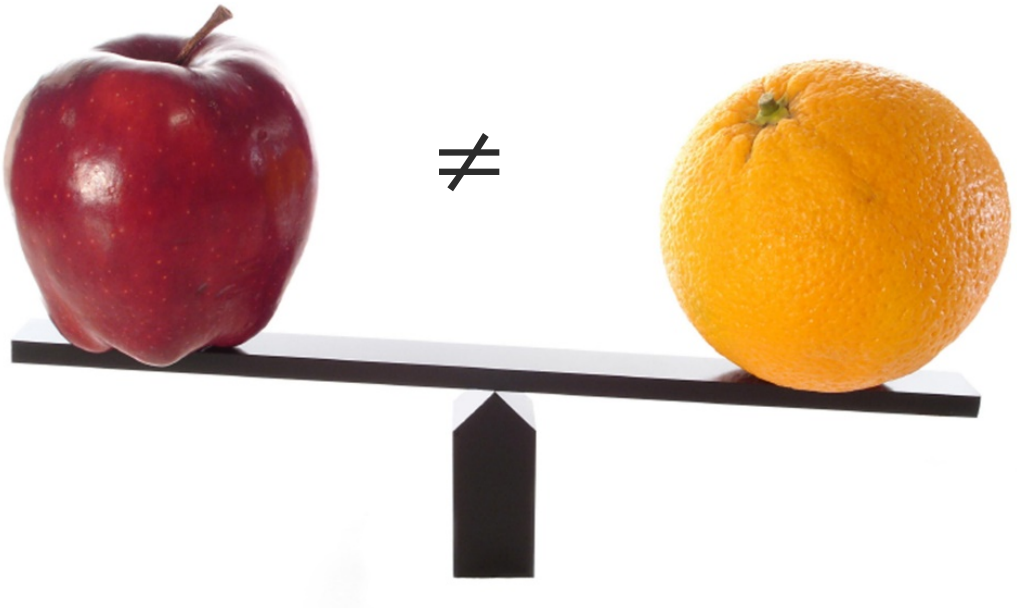
Process Mining

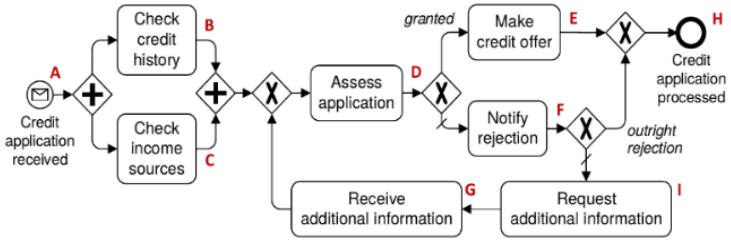


Conformance Checking

Given a process model and an event log, find, describe, and/or measure the *differences* between them

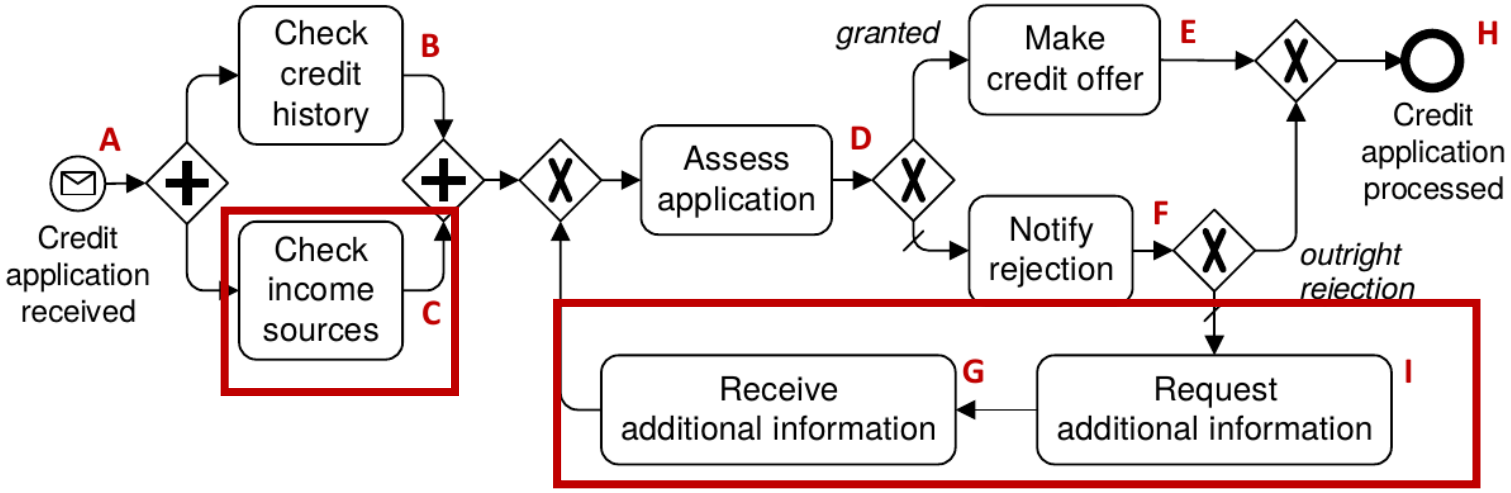
No. of Instances	Log Traces
1207	ABDEA
145	ACDGHFA
56	ACGDHFA
23	ACHDFA
28	ACDHFA





Conformance Checking with Trace Alignment

Conformance Checking in Apromore (Behavior Alignment)



Event log:

ABCDEH

ACBDEH

ABCDFH

ACBDFH

ABDEH

ABDFH

Unfitting behaviour:

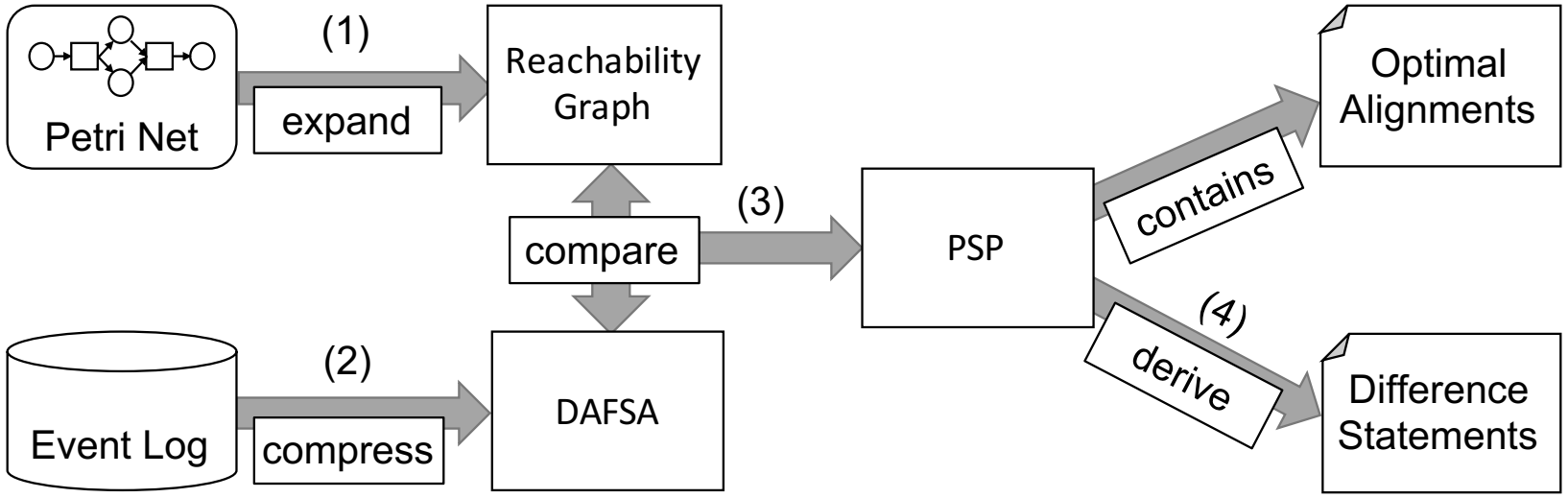
- Task C is *optional* (i.e. may be skipped) in the log

Additional behavior:

- The *cycle* including IGDF is not observed in the log

García-Bañuelos et al. "Complete and Interpretable Conformance Checking of Business Processes" *IEEE Transactions on Software Engineering* 44(3): 262-290, 2018

Behavior Alignment



Daniel Reißner, Raffaele Conforti, Marlon Dumas, Marcello La Rosa, Abel Armas-Cervantes: Scalable Conformance Checking of Business Processes. OTM Conferences (1) 2017: 607-627

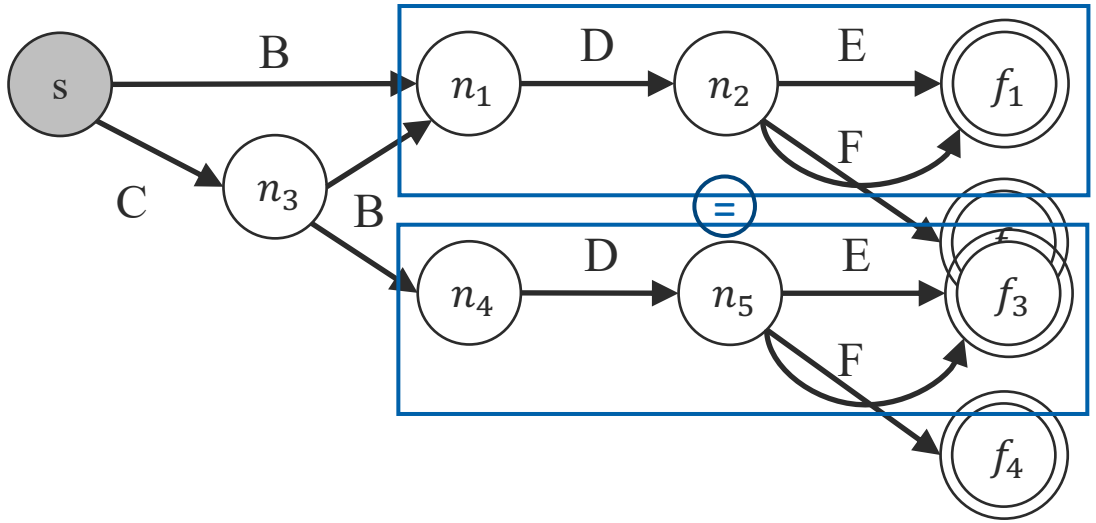
From event log to DAFSA

Prefixes	Suffixes
$\langle B, D \rangle, \langle C, B, D \rangle$	$\langle D, F \rangle, \langle D, E \rangle$

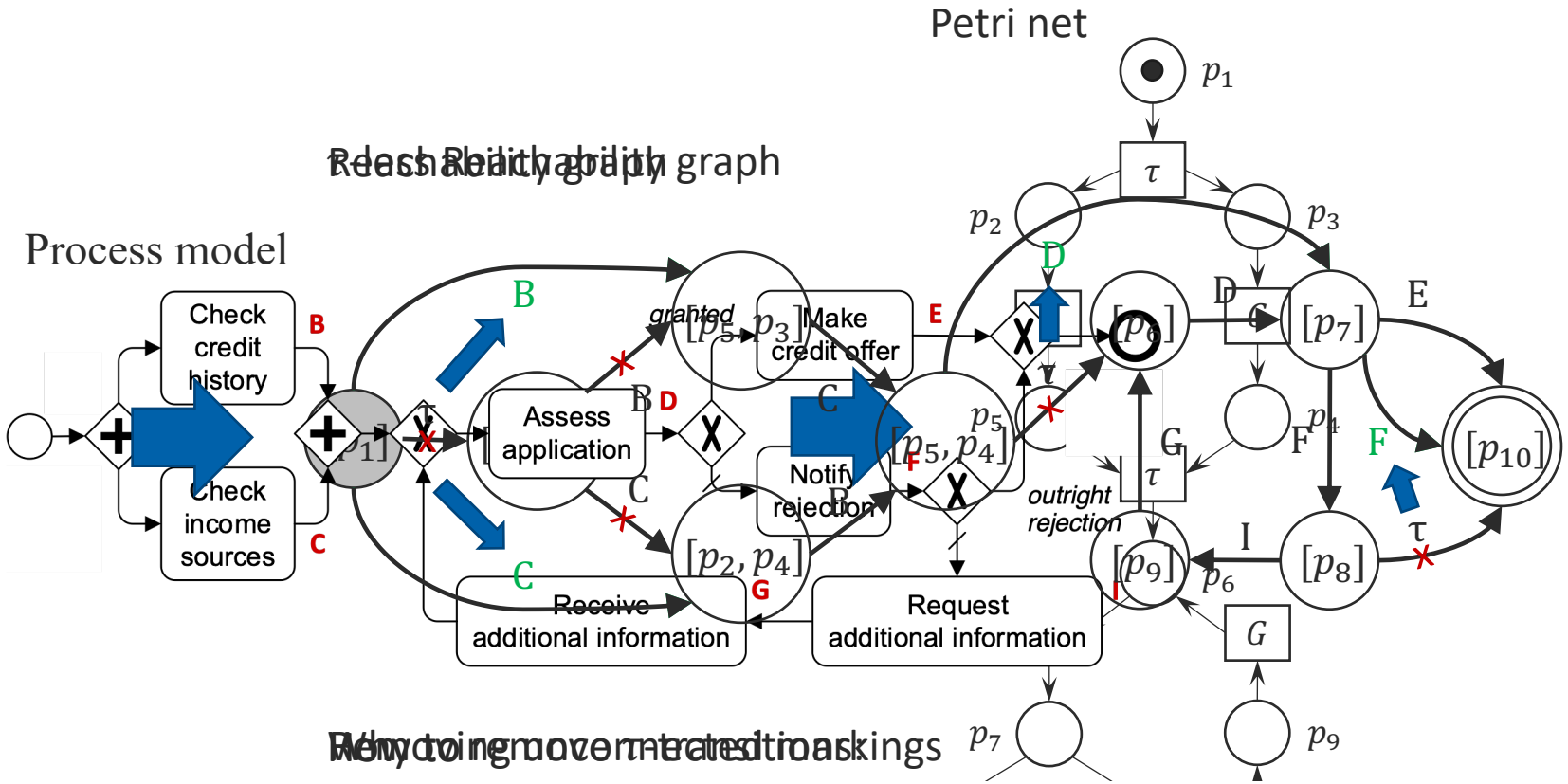
Log

Trace	N
$\langle B, D, E \rangle$	5
$\langle B, D, F \rangle$	10
$\langle C, B, D, E \rangle$	15
$\langle C, B, D, F \rangle$	5

DAFSA

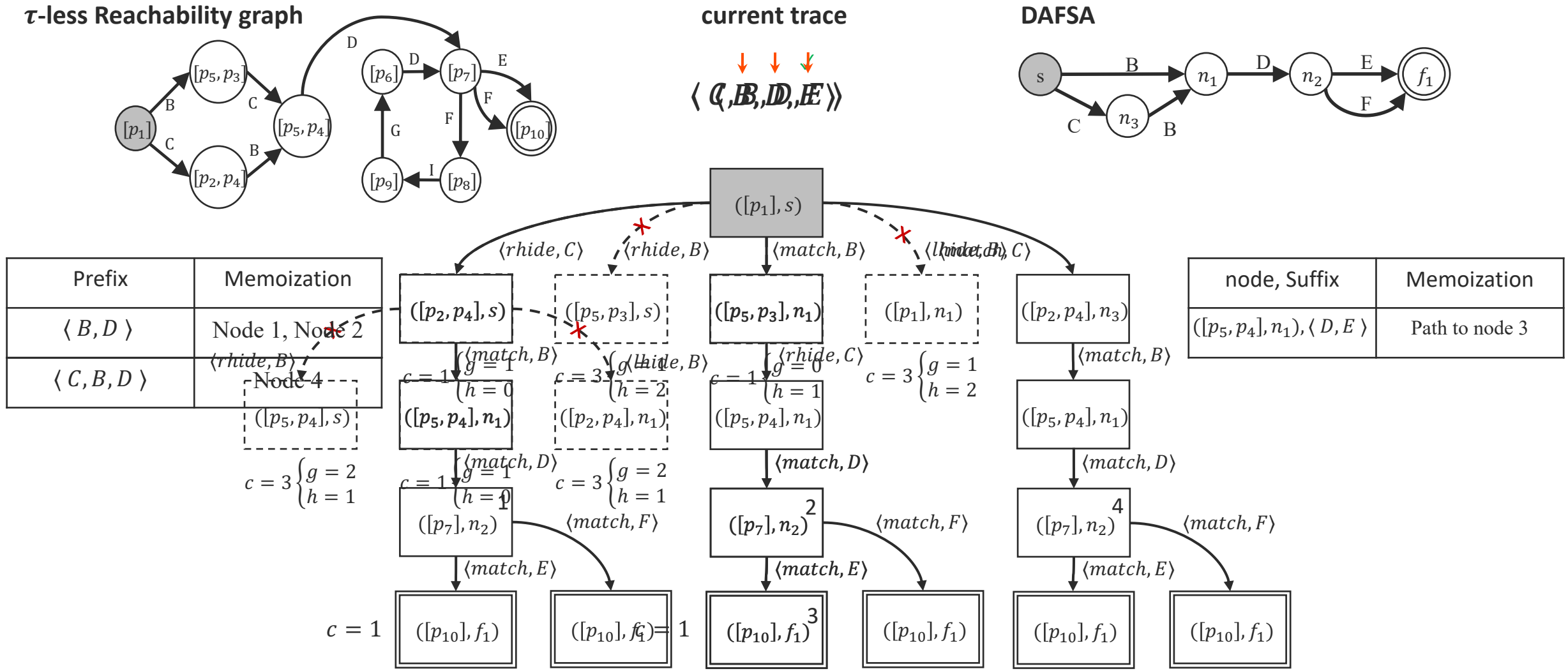


From process model to reachability graph

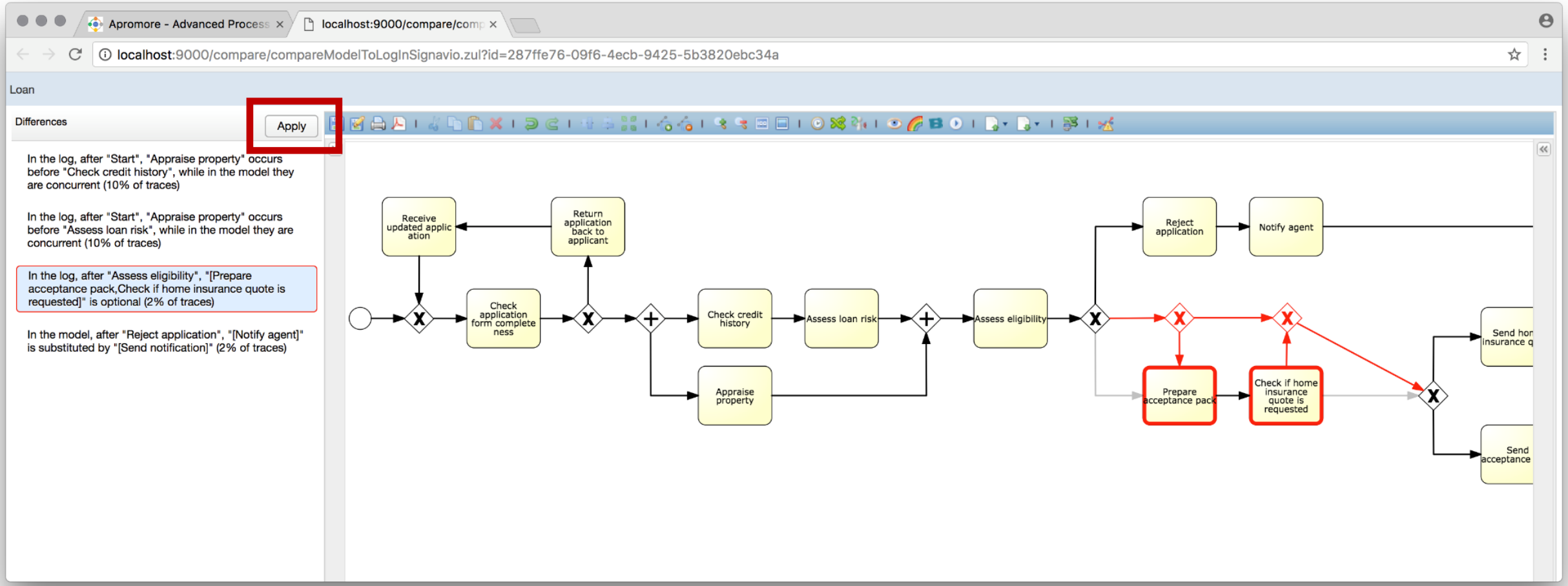


- For each τ not targeting a final marking, insert a copy of each outgoing arc of the target of τ and link it to the source,
- otherwise, use each incoming arc of its source

PSP construction with the A*- Algorithm



Interactive Model Repair

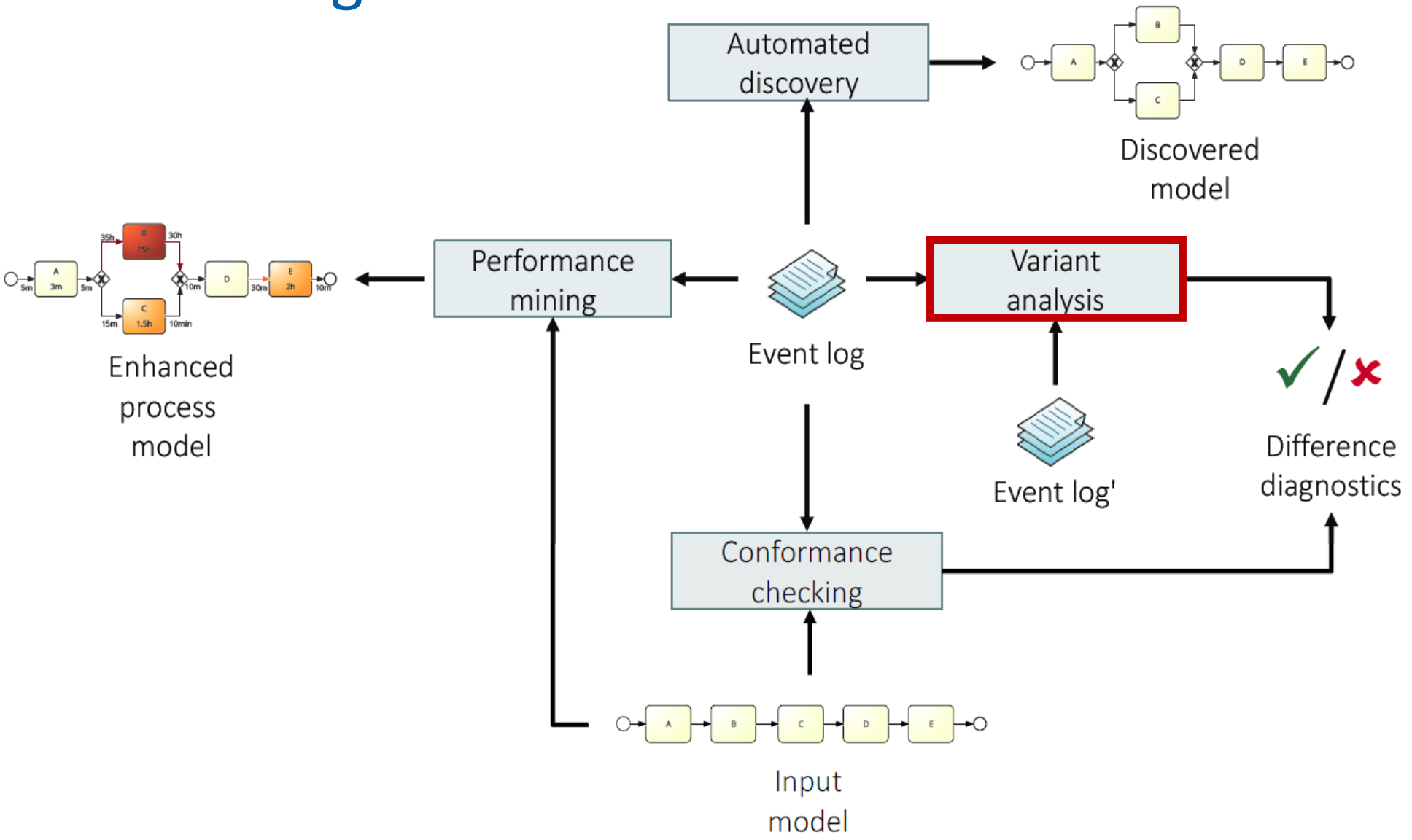


A. Armas Cervantes et al. “Interactive and Incremental Business Process Model Repair”, *Proceedings of CoopIS*’2017

UNIVERSITY OF TARTU

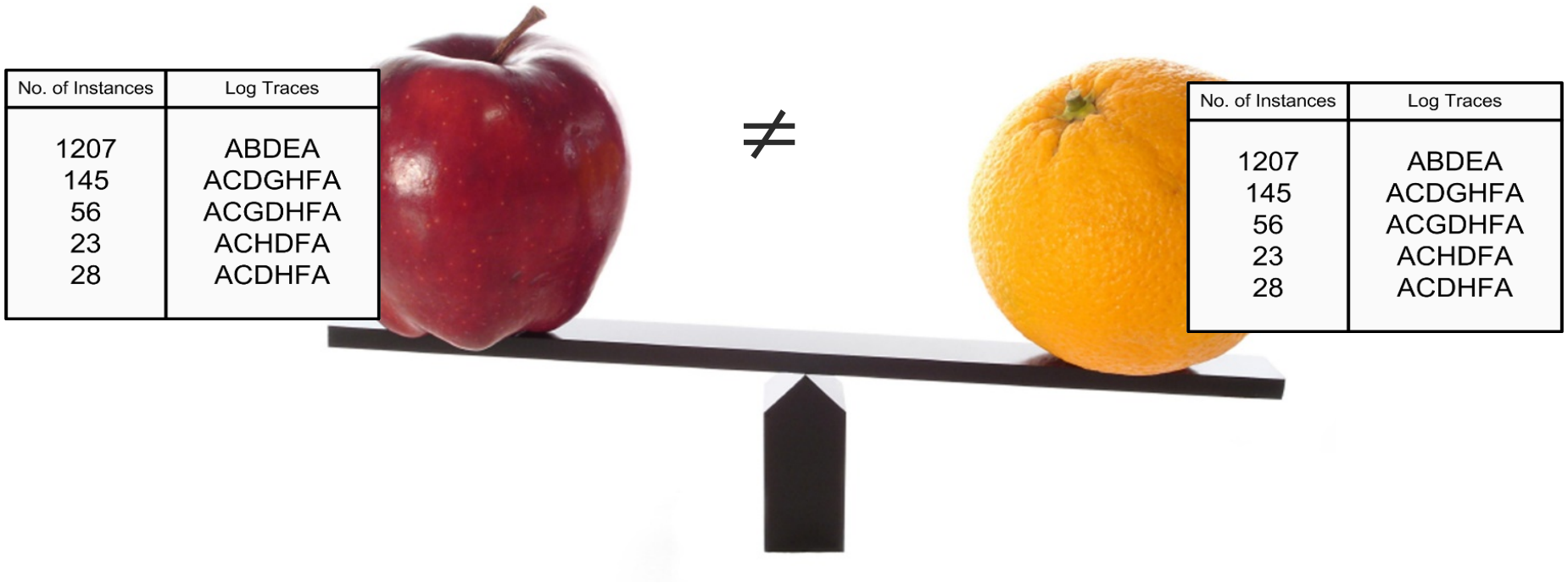
Demo Time!

Process Mining



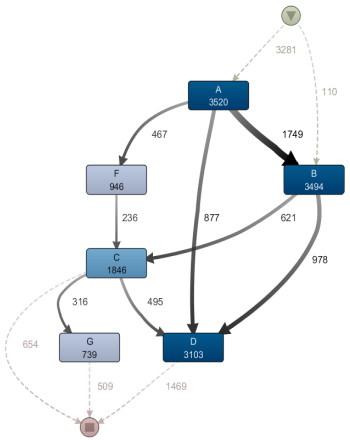
Variants Analysis

Given two logs, find the *differences* and *root causes* for variation between the two logs

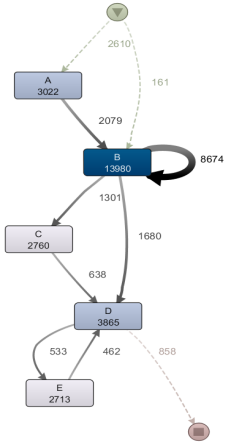


Variants Analysis

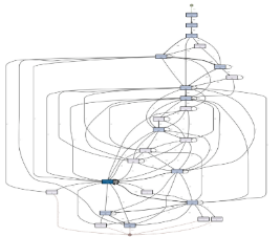
- Model comparison



Model Comparison



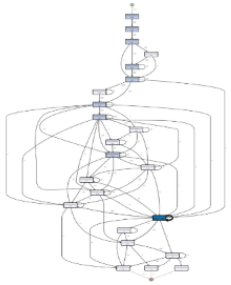
- Log delta analysis



L1 - Short stay
448 cases
7329 events

Log Delta Analysis

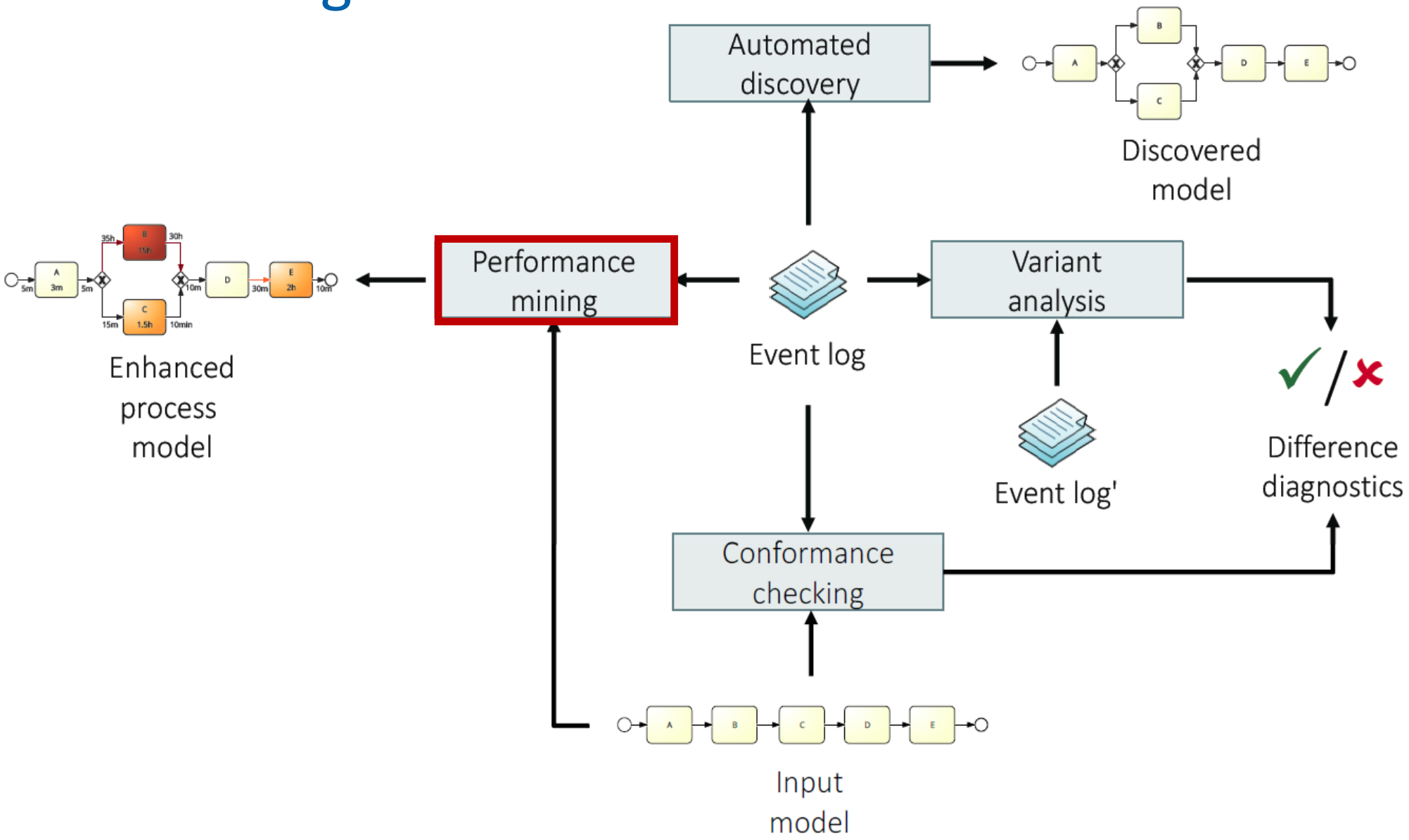
In L1, "Nursing Primary Assessment" is repeated after "Medical Assign" and "Triage Request", while in L2 it is not…



L2 - Long stay
363 cases
7496 events

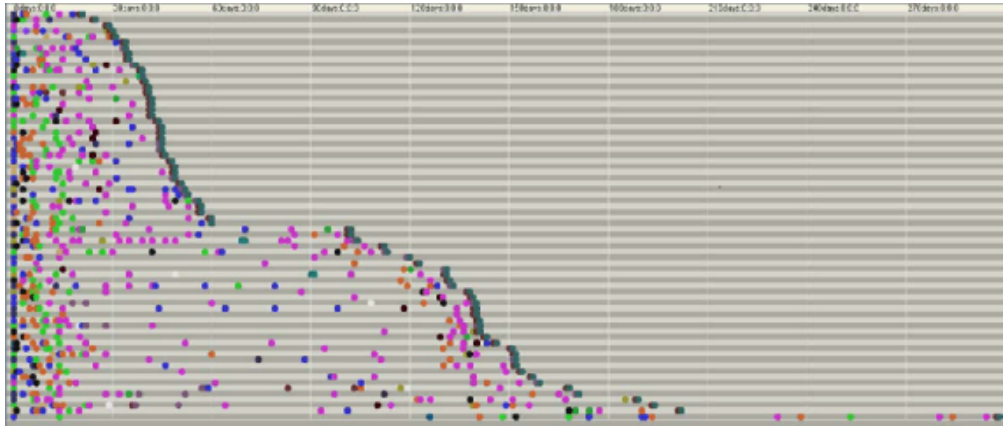
N. van Beest et al. "Log Delta Analysis: Interpretable Differencing of Business Process Event Logs" Proc. of BPM'2015

Process Mining

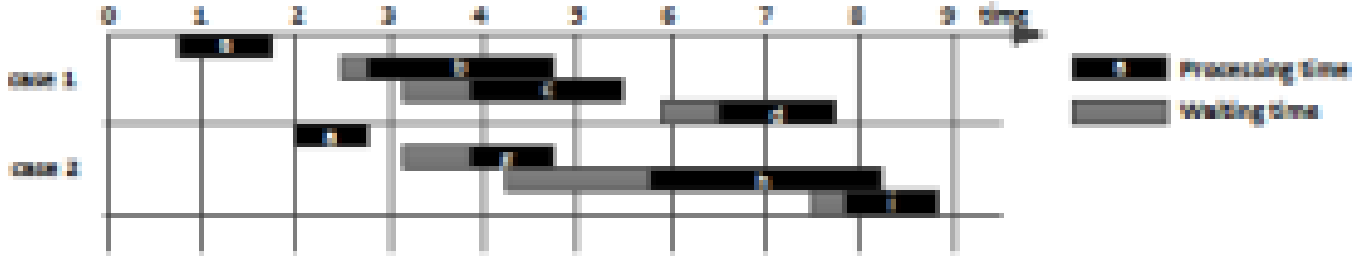


Performance Mining

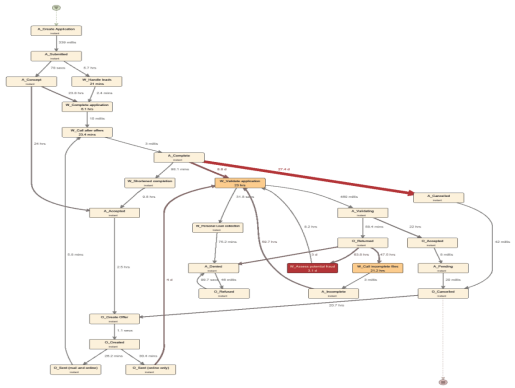
- Dotted charts



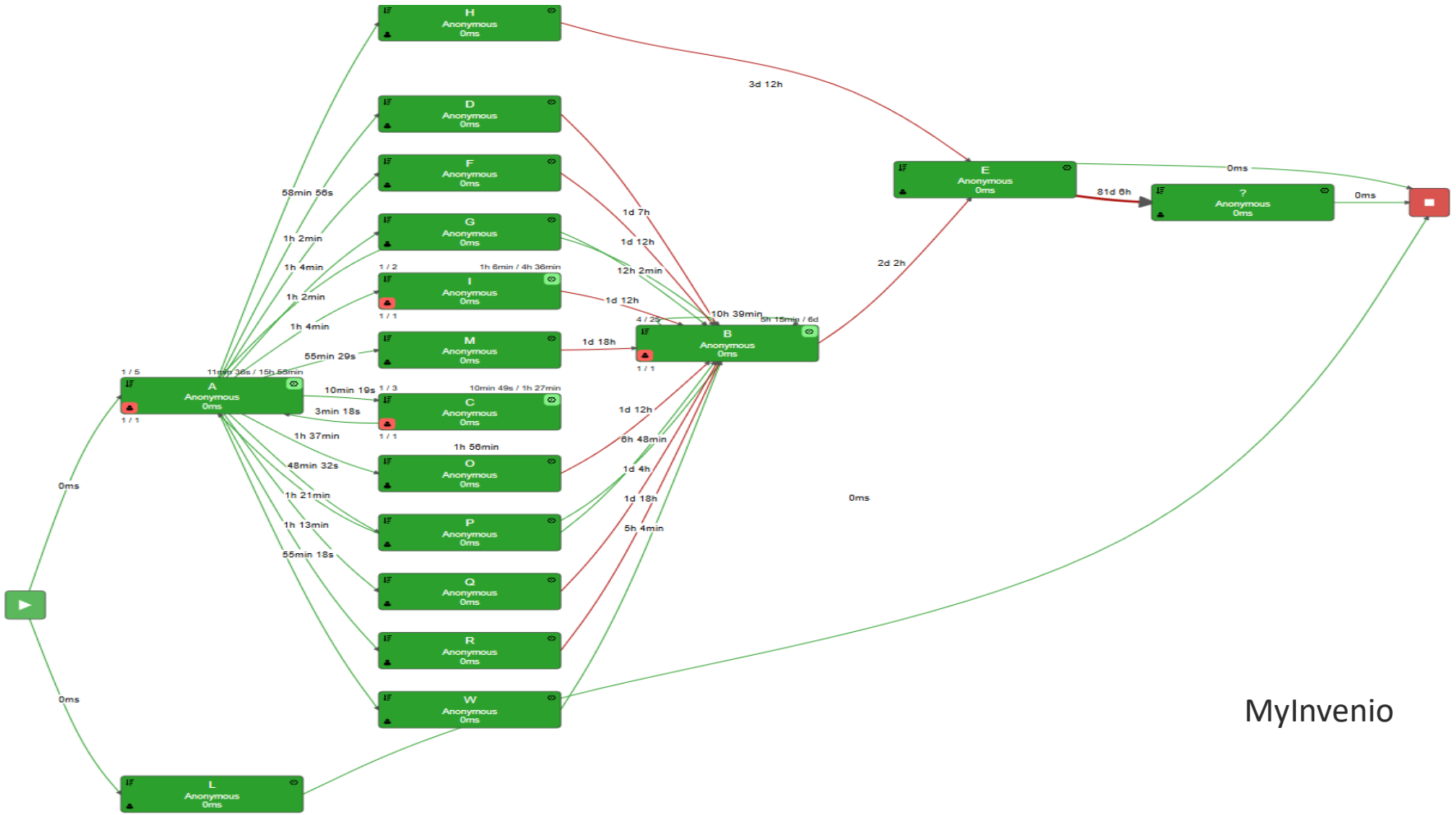
- Timeline diagrams



- Performance-enhanced dependency graphs

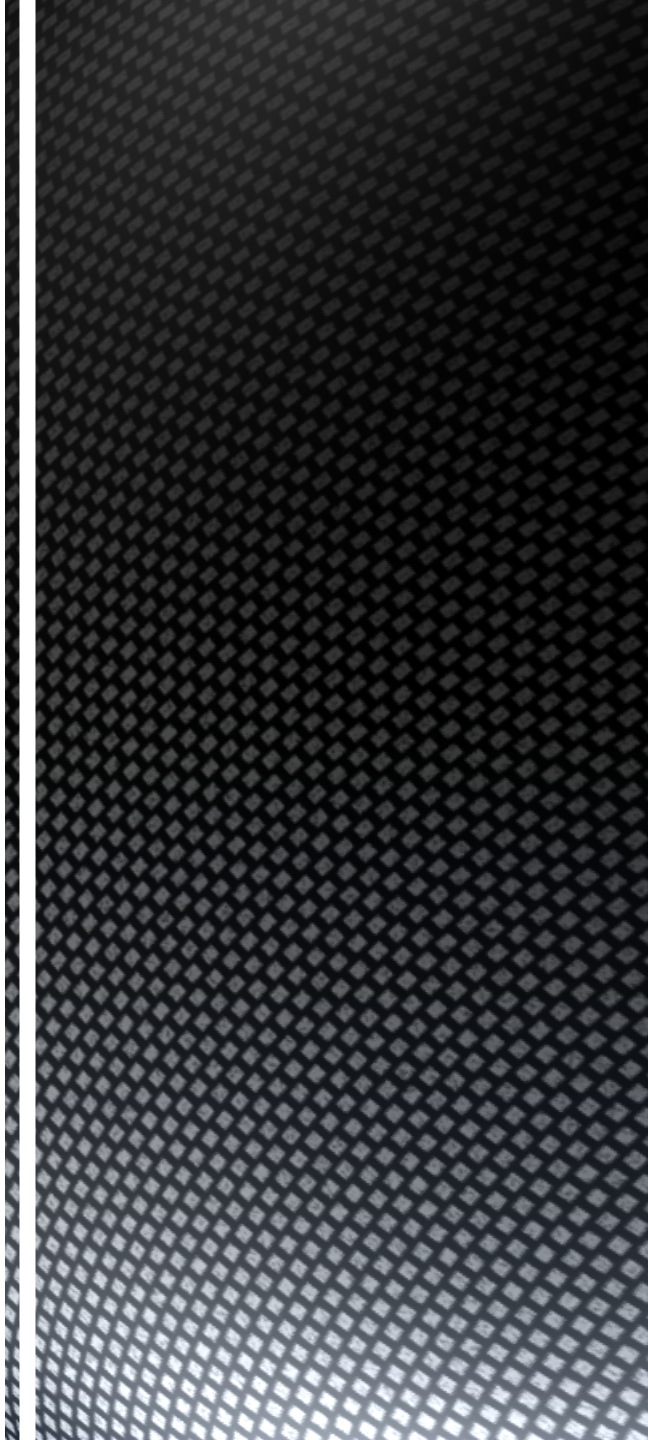


Performance-Enhanced Handoff Map

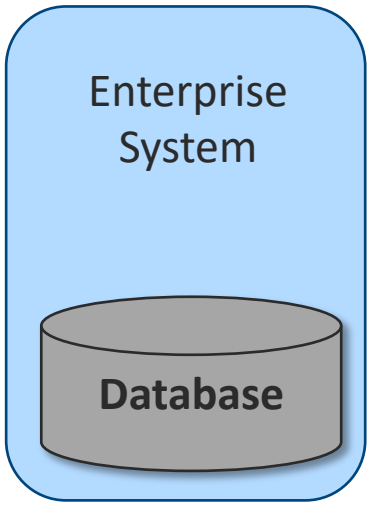


MyInvenio

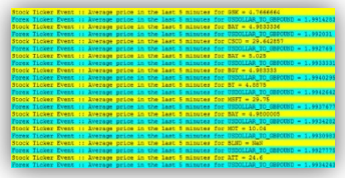
Operational Process Analytics



Business Process Analytics

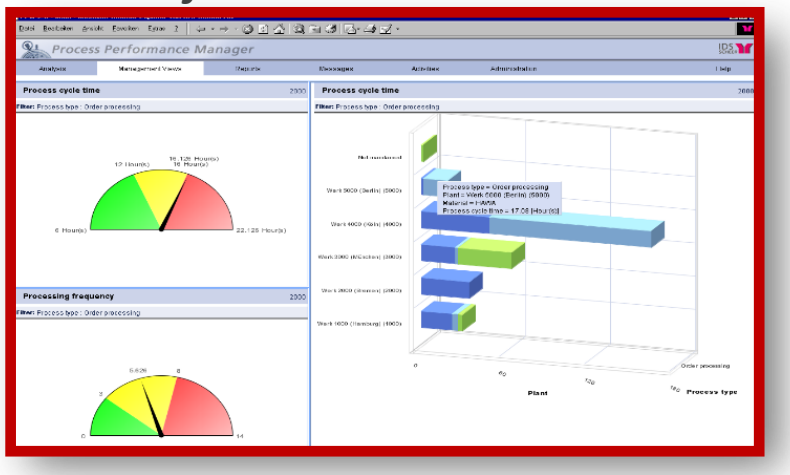


Event stream

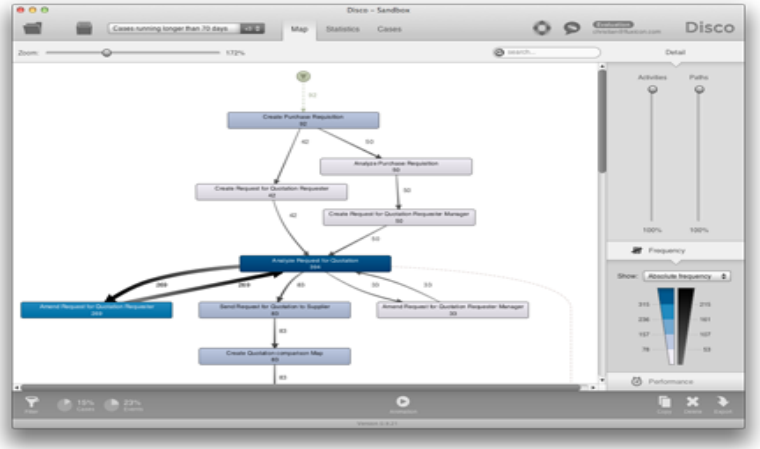


Event log

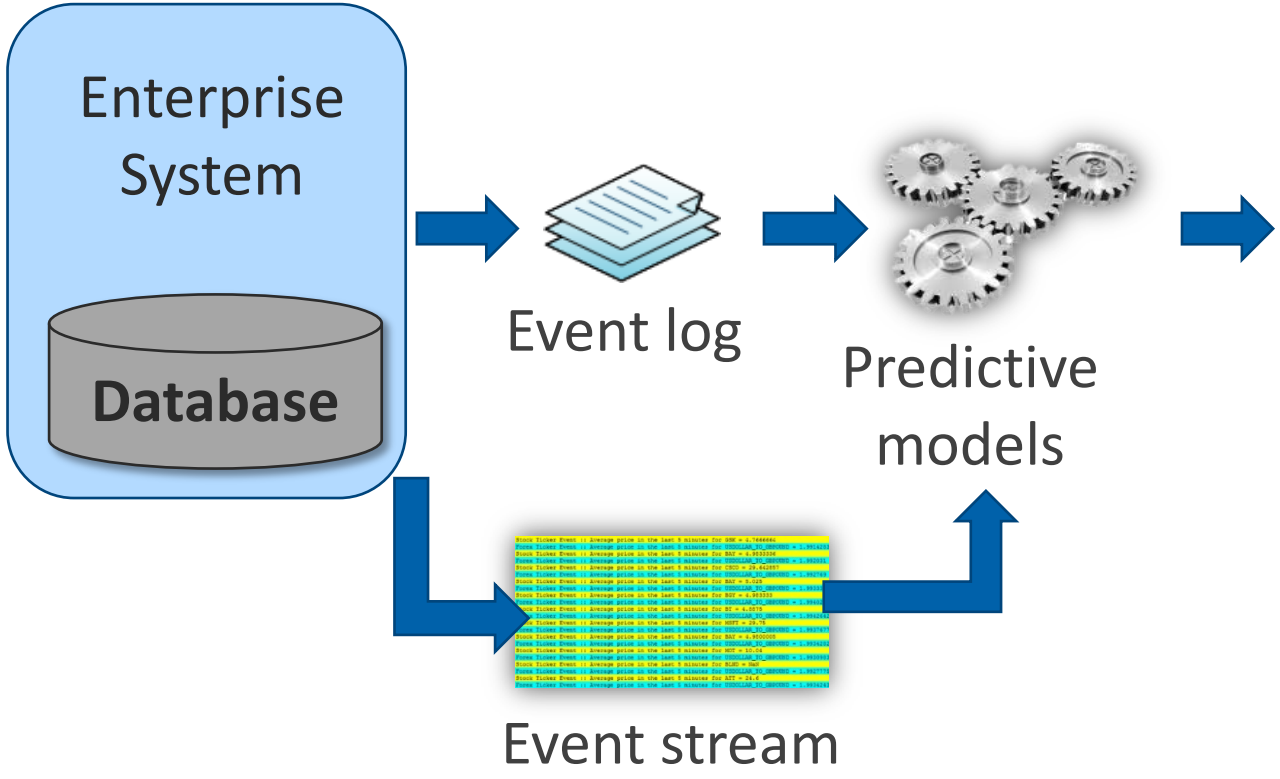
Statistics-Based Techniques
Performance Dashboards



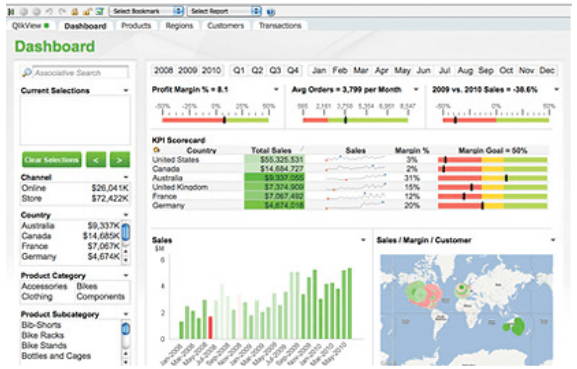
Model-Based Techniques
Process Mining



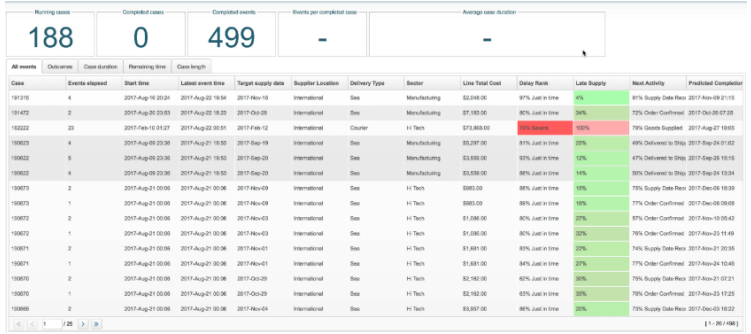
Predictive Process Monitoring



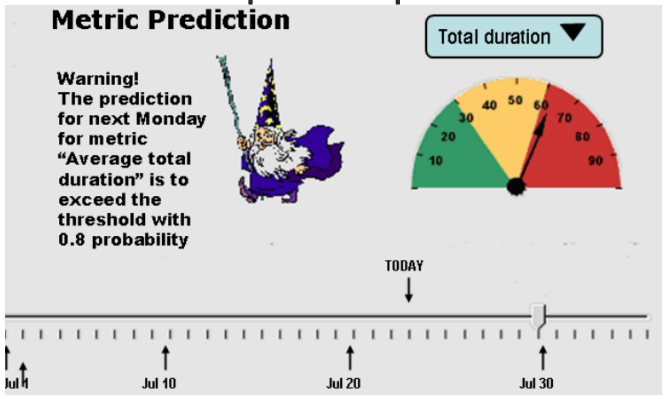
Aggregate predictive dashboards



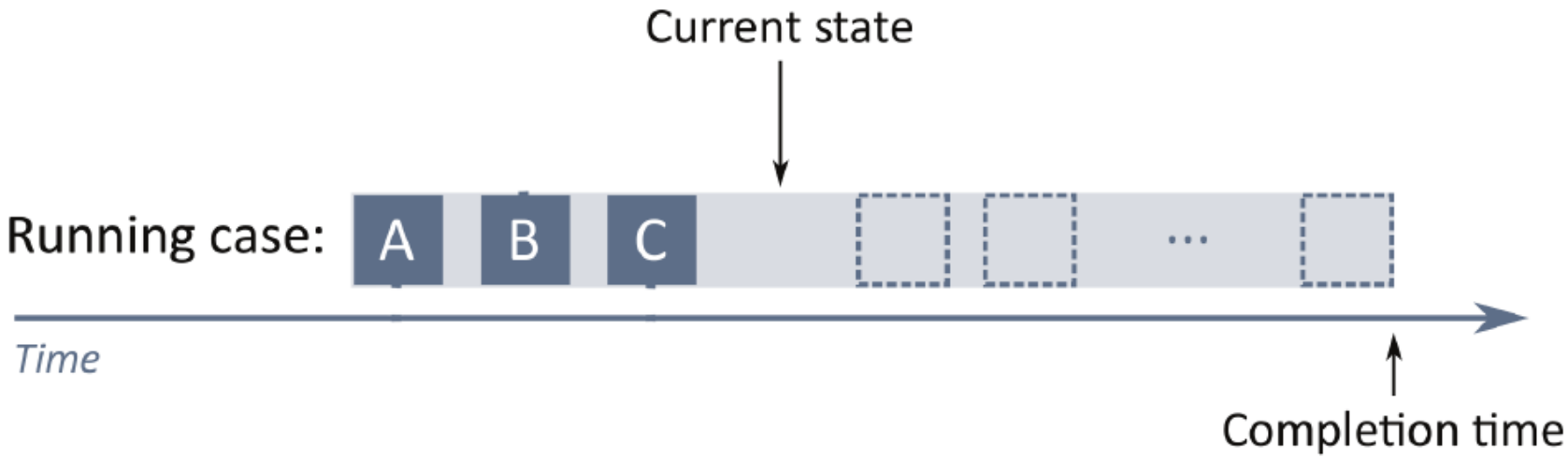
Detailed predictive dashboard



Alarm-based prescriptive dashboard

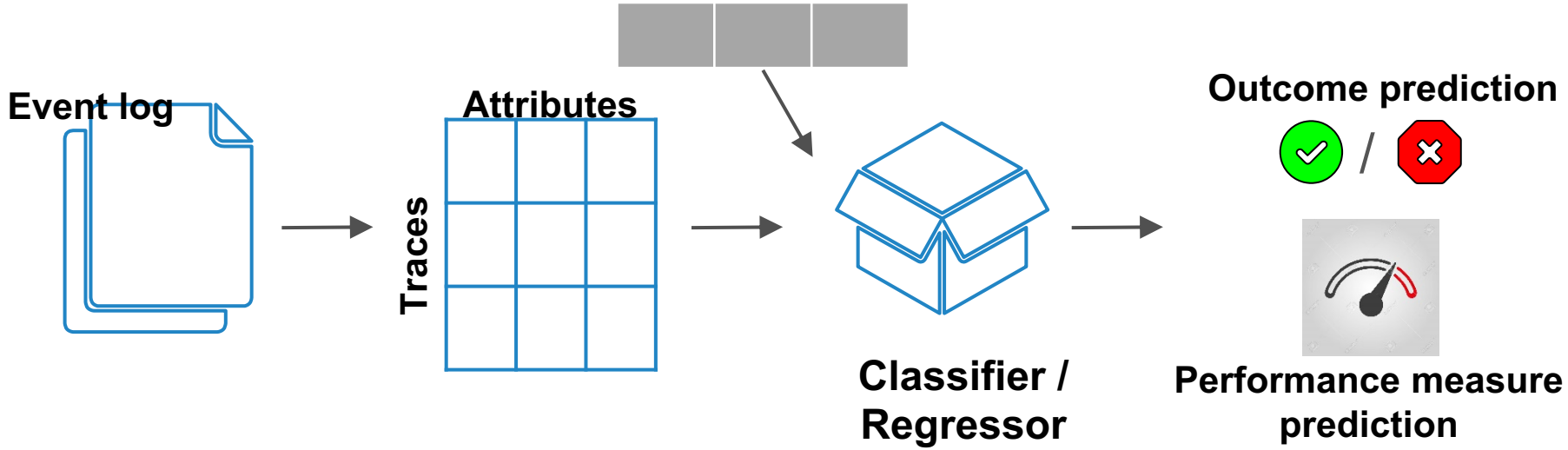


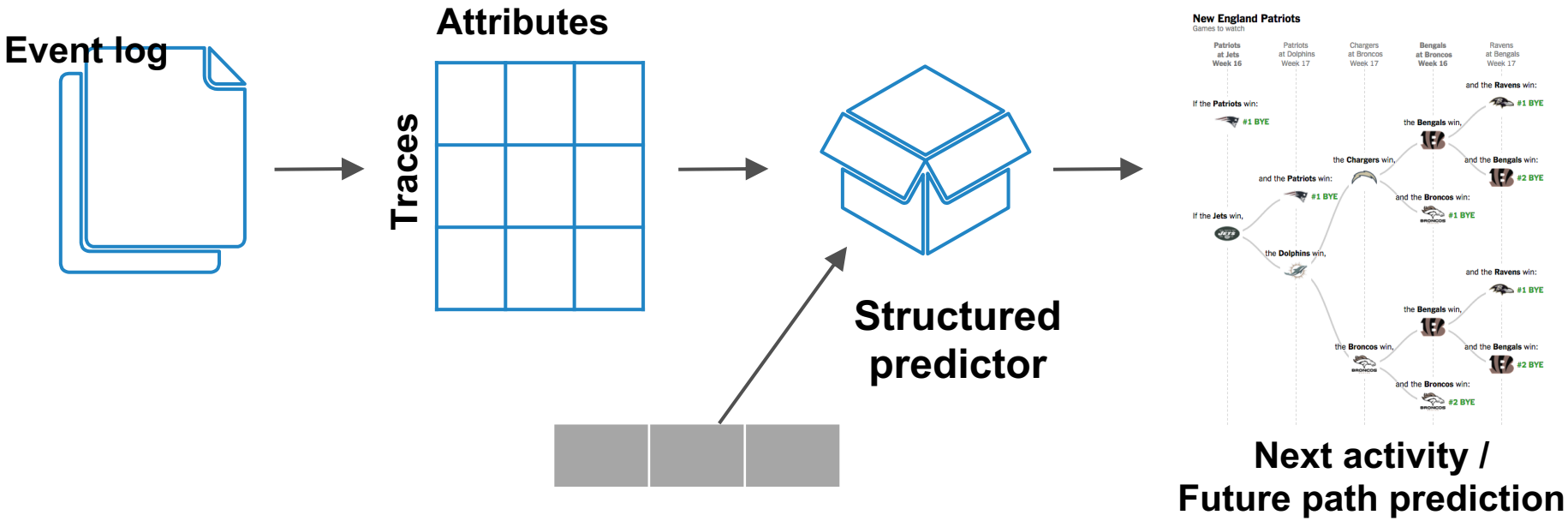
Predictive Process Monitoring



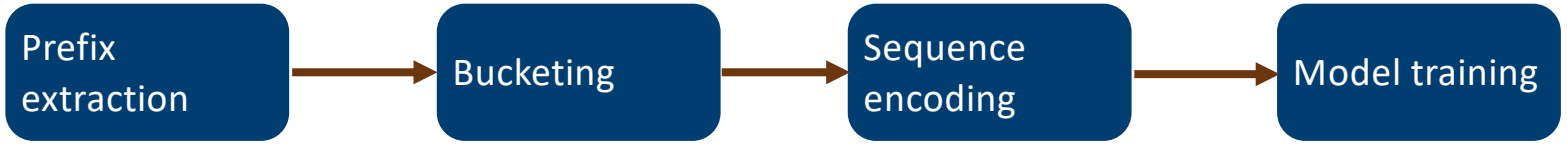
- What is the next activity for this case?
- When is this next activity going to take place?
- How long is this case still going to take until it is finished?
- What is the outcome of this case?
- Is the compensation going to be paid? Or rejected?

Predictive Process Monitoring: General Approach

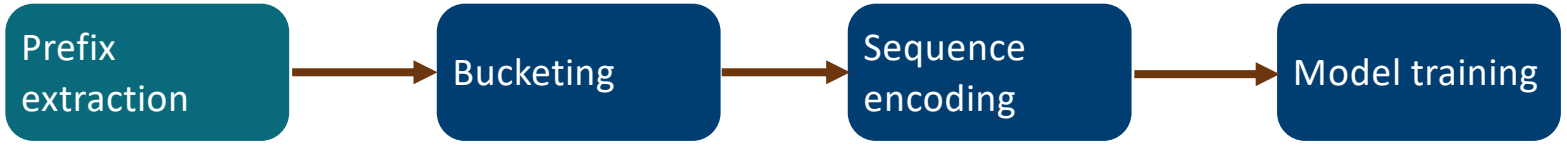


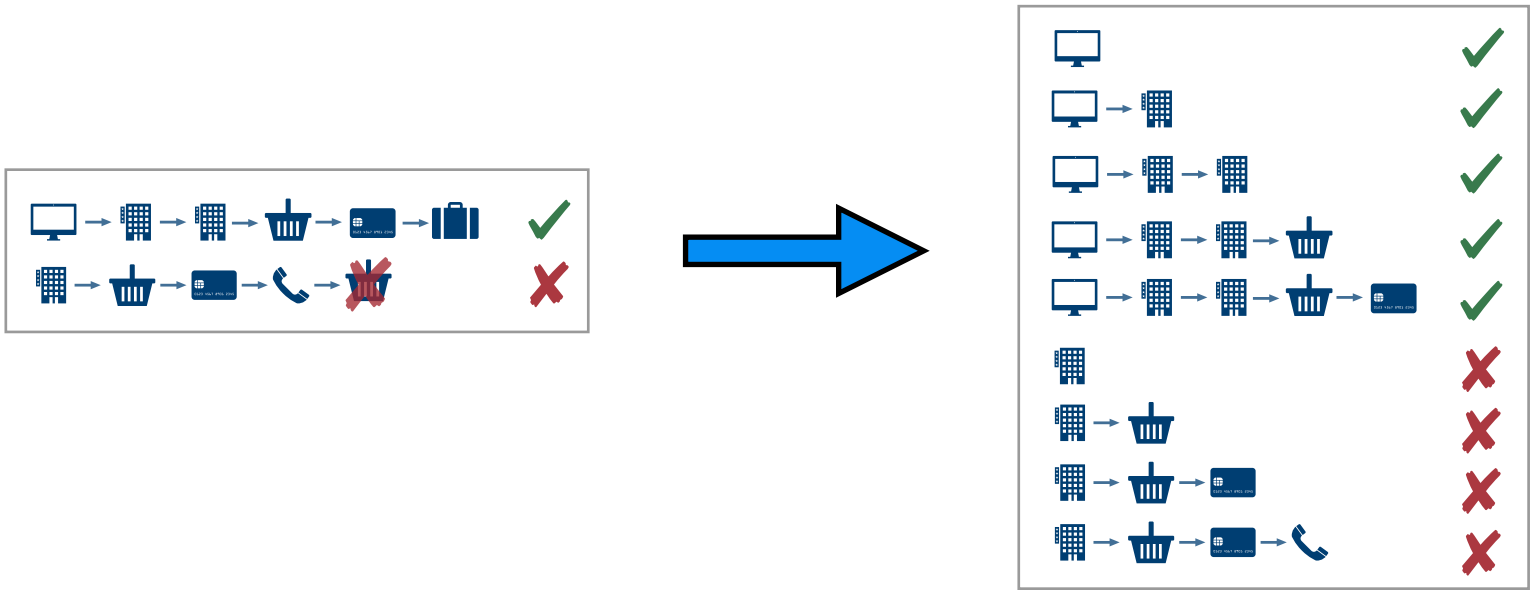


Predictive process monitoring workflow

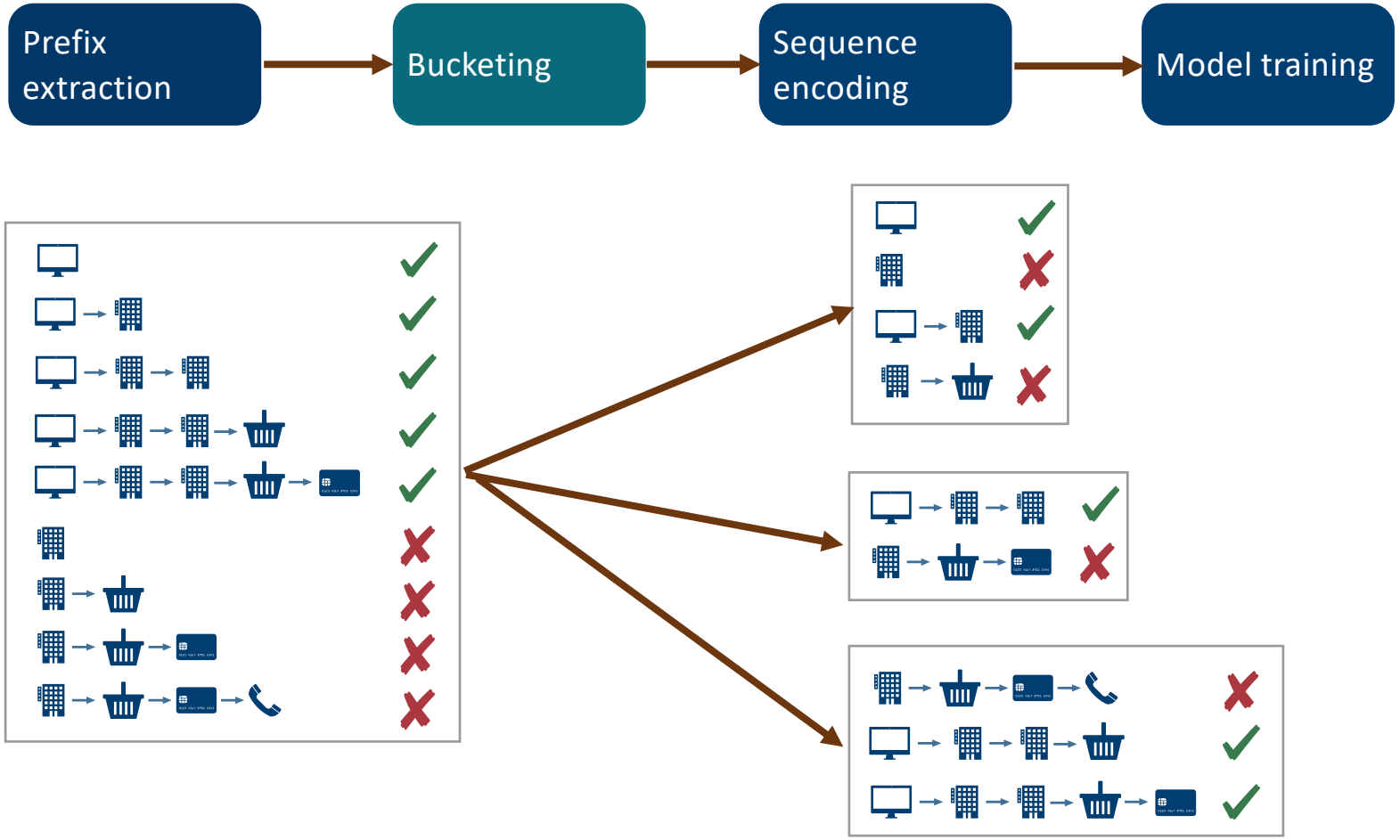


Prefix extraction

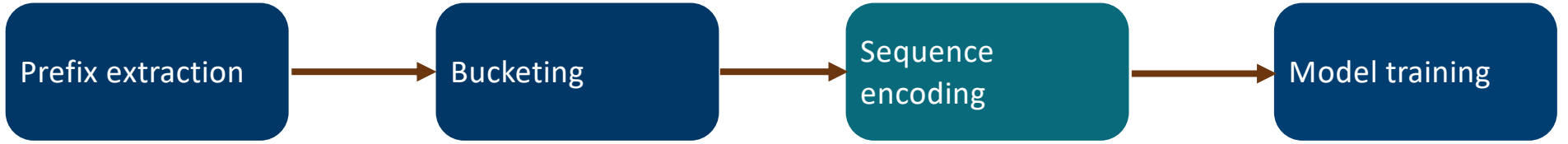


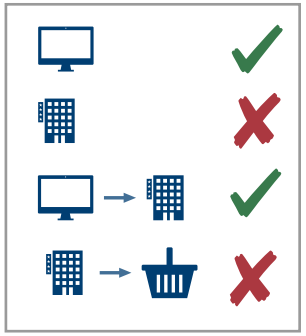


Bucketing

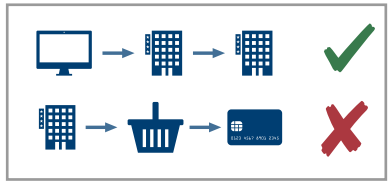

Prefix extraction
Bucketing
Sequence encoding
Model training

Sequence encoding

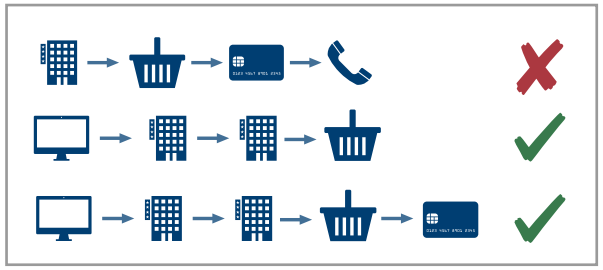




(Amsterdam,	2 adults,	€100)	not cancelled
(Paris,	1 adult,	€150)	cancelled
(Amsterdam,	2 adults,	€100)	not cancelled
(Paris,	1 adult,	€150)	cancelled



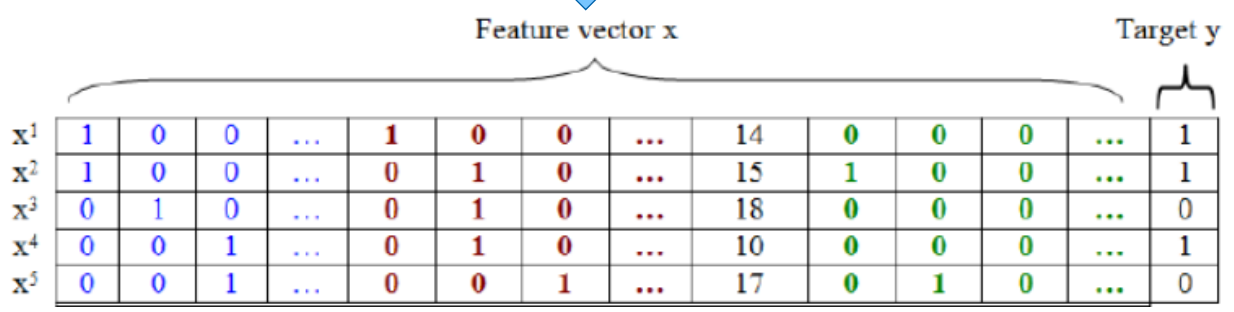
(Amsterdam,	2 adults,	3 days,	€100)	not cancelled
(Paris,	1 adult,	4 days,	€150)	cancelled



(Amsterdam,	2 adults,	3 days,	July,	€100)	not cancelled
(Amsterdam,	2 adults,	3 days,	April,	€100)	not cancelled
(Paris,	1 adult,	4 days,	July,	€150)	cancelled

Sequence encoding

Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...	...	...	...	...	...	...



Sequence encoding

- Index-based encoding

id	loan goal	activity_1	...	activity_m	amount_1	...	amount_m	label
1	mortgage	check completeness	...	make offer	0	...	80 000	1
2	car	check completeness	...	calculate risk score	0	...	0	0

Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...	...	...	...	...	...	...

Index-based encoding + prefix-length bucketing

Case ID	Time-stamp	Activity	Resource	Amount
C1	T1	A	R1	100
C1	T2	B	R1	100
C1	T3	D	R2	100
C1	T4	E	R3	100
C1	T5	E	R3	100
C2	T1	A	R1	15
C2	T2	B	R1	50
C2	T3	D	R4	50
C2	T4	F	R4	50
C3	T1	A	R5	30
C3	T2	C	R5	30
C3	T3	B	R5	30

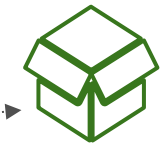
Activity1	Resource1	Amount1
A	R1	100
A	R1	15
A	R5	30

Classifier 1

Index-based encoding + prefix-length bucketing

Case ID	Time-stamp	Activity	Resource	Amount
C1	T1	A	R1	100
C1	T2	B	R1	100
C1	T3	D	R2	100
C1	T4	E	R3	100
C1	T5	E	R3	100
C2	T1	A	R1	15
C2	T2	B	R1	50
C2	T3	D	R4	50
C2	T4	F	R4	50
C3	T1	A	R5	30
C3	T2	C	R5	30
C3	T3	B	R5	30

Activity1	Resource1	Amount1
A	R1	100
A	R1	15
A	R5	30



Classifier 1

Activity1	Resource1	Amount1	Activity2	Resource2	Amount2
A	R1	100	B	R1	100
A	R1	15	B	R1	50
A	R5	30	C	R5	30

Classifier 2

Sequence encoding

- Index-based encoding

id	loan goal	activity_1	...	activity_m	amount_1	...	amount_m	label
1	mortgage	check completeness	...	make offer	0	...	80 000	1
2	car	check completeness	...	calculate risk score	0	...	0	0

- Aggregation encoding

id	loan goal	count(check compl-ness)	...	count(make offer)	min(amt)	...	max(amt)	label
1	mortgage	1	...	1	0	...	80 000	1
2	car	1	...	0	0	...	0	0

Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...	...	...	...	...	...	...

Aggregation encoding

Case ID	Time-stamp	Activity	Resource	Amount
C1	T1	A	R1	100
C1	T2	B	R1	100
C1	T3	D	R2	100
C1	T4	E	R3	100
C1	T5	E	R3	100
C2	T1	A	R1	15
C2	T2	B	R1	50
C2	T3	D	R4	50
C2	T4	F	R4	50
C3	T1	A	R5	30
C3	T2	C	R5	30
C3	T3	B	R5	30

Time-stamp	Freq A	Freq B	Freq C	Freq R1	Freq R5	Max Amount	Mean Amount
T1	1	0	0	1	0	100	100
T1	1	0	0	1	0	15	15
T1	1	0	0	0	1	30	30

Aggregation encoding

Case ID	Time-stamp	Activity	Resource	Amount
C1	T1	A	R1	100
C1	T2	B	R1	100
C1	T3	D	R2	100
C1	T4	E	R3	100
C1	T5	E	R3	100
C2	T1	A	R1	15
C2	T2	B	R1	50
C2	T3	D	R4	50
C2	T4	F	R4	50
C3	T1	A	R5	30
C3	T2	C	R5	30
C3	T3	B	R5	30

Time-stamp	Freq A	Freq B	Freq C	Freq R1	Freq R5	Max Amount	Mean Amount
T1	1	0	0	1	0	100	100
T1	1	0	0	1	0	15	15
T1	1	0	0	0	1	30	30
T2	1	1	0	2	0	100	100
T2	1	1	0	2	0	50	32.5
T2	1	0	1	0	2	30	30

Classifier

Sequence encoding

Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...	...	...	...	...	...	...

- Index-based encoding

id	loan goal	activity_1	...	activity_m	amount_1	...	amount_m	label
1	mortgage	check completeness	...	make offer	0	...	80 000	1
2	car	check completeness	...	calculate risk score	0	...	0	0

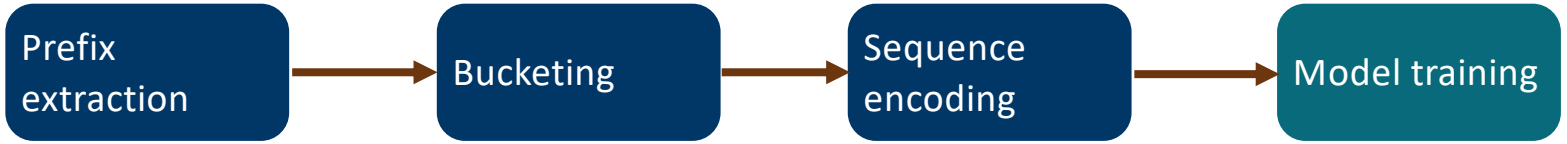
- Aggregation encoding

id	loan goal	count(check compl-ness)	...	count(make offer)	min(amt)	...	max(amt)	label
1	mortgage	1	...	1	0	...	80 000	1
2	car	1	...	0	0	...	0	0

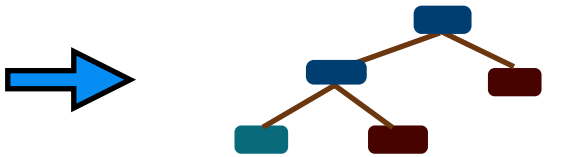
- LSTM

id	event	loan goal	event	amount	label
1	1	mortgage	check completeness	0	1
	...	...	...	...	
	m	mortgage	make offer	80 000	
2	1	car	check completeness	0	0
	...	...	...	...	
	m	car	calculate risk score	0	

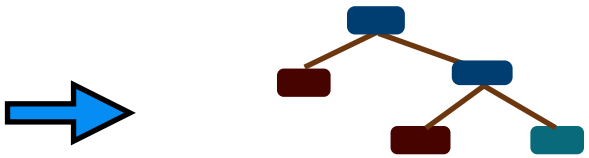
Model training



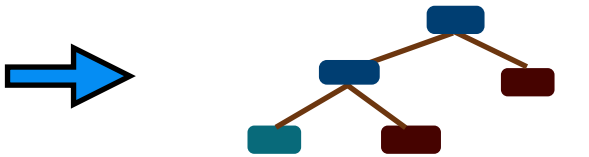
(Amsterdam,	2 adults,	€100)	not cancelled
(Paris,	1 adult,	€150)	cancelled
(Amsterdam,	2 adults,	€100)	not cancelled
(Paris,	1 adult,	€150)	cancelled



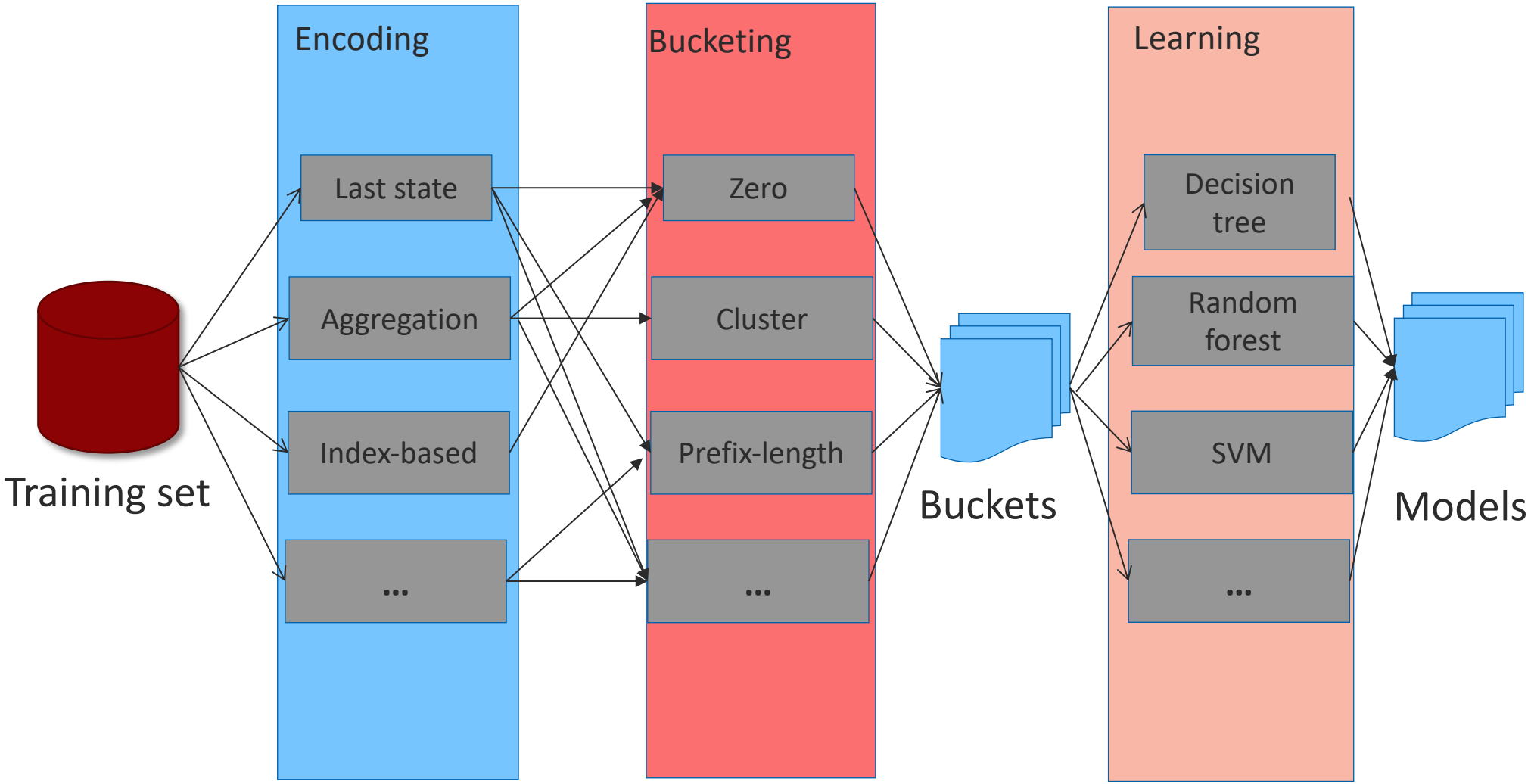
(Amsterdam,	2 adults,	3 days,	€100)	not cancelled
(Paris,	1 adult,	4 days,	€150)	cancelled



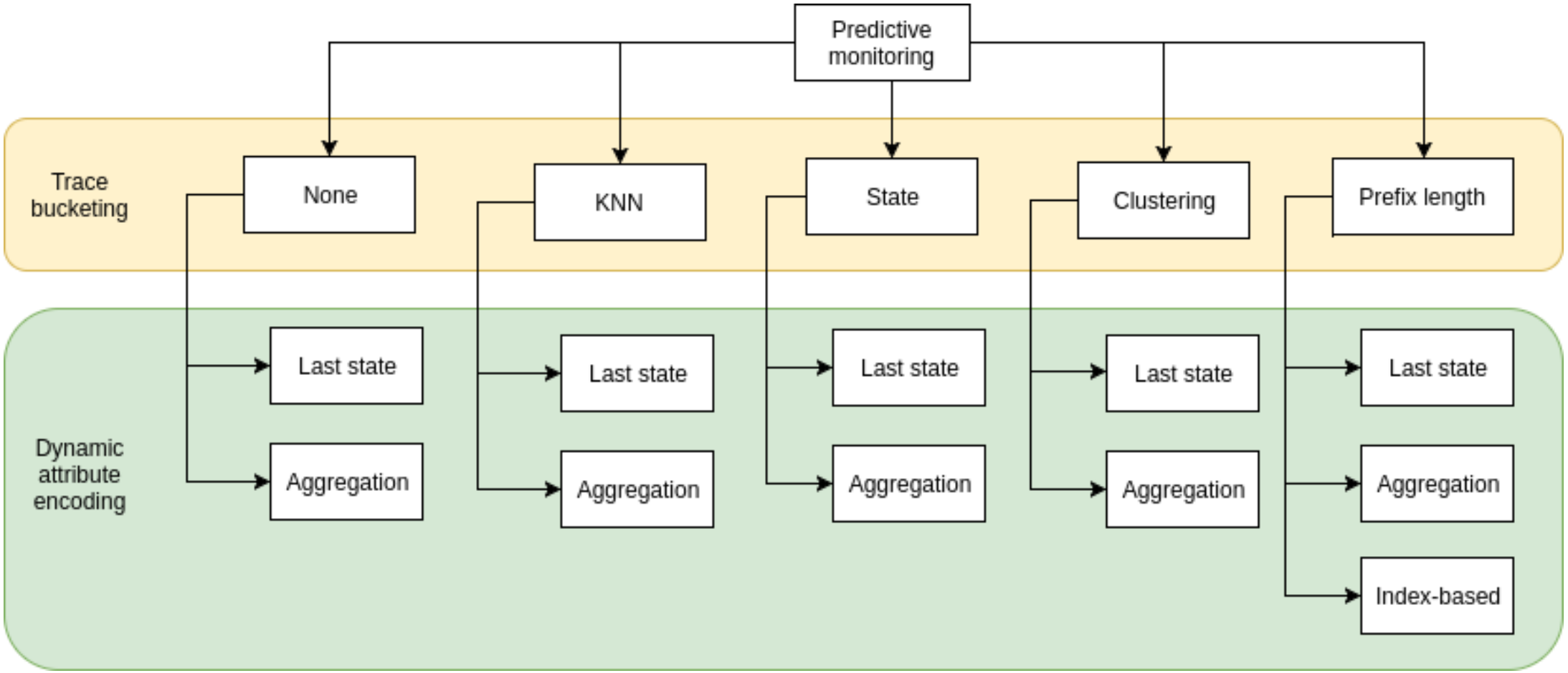
(Amsterdam,	2 adults,	3 days,	July,	€100)	not cancelled
(Amsterdam,	2 adults,	3 days,	April,	€100)	not cancelled
(Paris,	1 adult,	4 days,	July,	€150)	cancelled



Predictive process monitoring workflow



Taxonomy of existing approaches



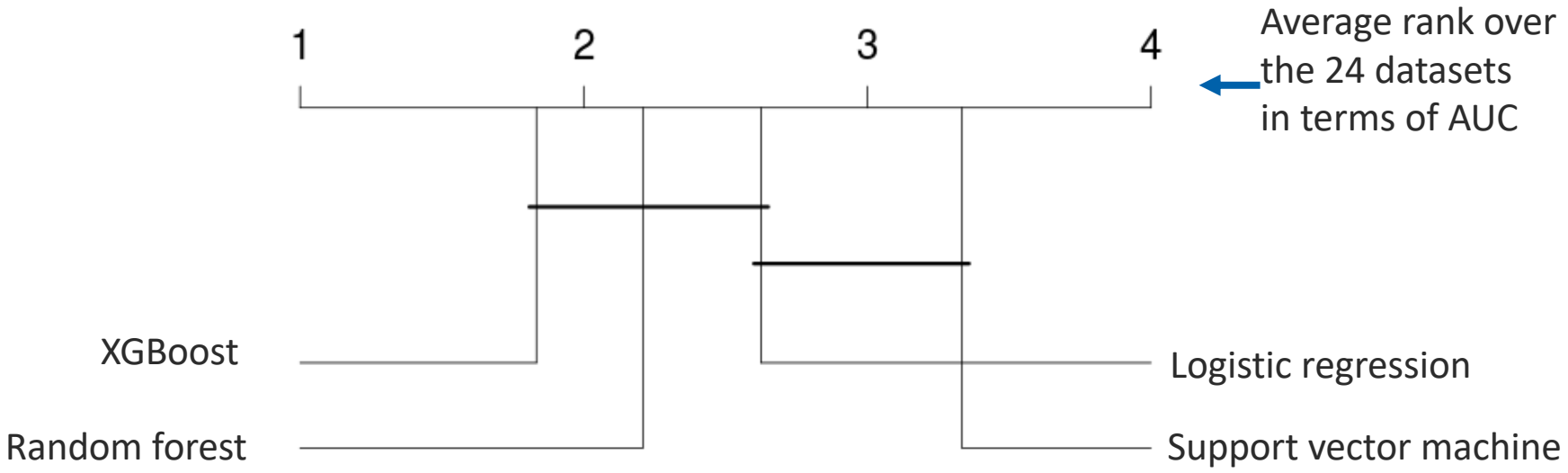
What is the relative performance of these methods?

Irene Teinemaa, Marlon Dumas, Marcello La Rosa, Fabrizio Maria Maggi:
Outcome-Oriented Predictive Process Monitoring: Review and Benchmark. TKDD 13(2): 17:1-17:57 (2019)

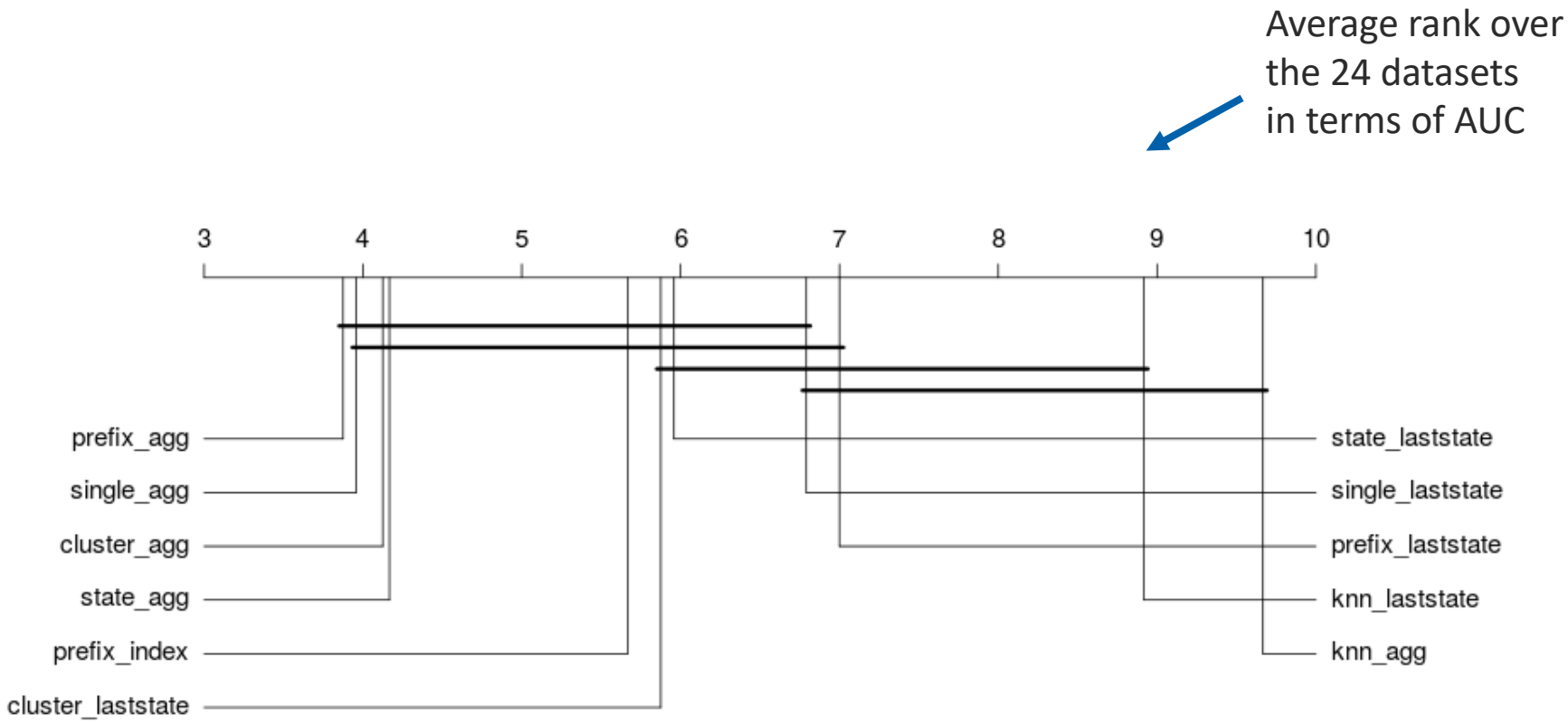
Evaluation datasets

	Dataset	Domain	# traces	Median # events in trace	Class ratio
1	bpic2011_1	Hospital treatment	1140	25	0.4
2	bpic2011_2	Hospital treatment	1140	54.5	0.78
3	bpic2011_3	Hospital treatment	1121	21	0.23
4	bpic2011_4	Hospital treatment	1140	44	0.28
5	bpic2012_1	Loan application	4685	35	0.48
6	bpic2012_2	Loan application	4685	35	0.17
7	bpic2012_3	Loan application	4685	35	0.35
8	bpic2015_1	Building permit	696	42	0.23
9	bpic2015_2	Building permit	753	55	0.19
10	bpic2015_3	Building permit	1328	42	0.2
11	bpic2015_4	Building permit	577	42	0.16
12	bpic2015_5	Building permit	1051	50	0.31
13	bpic2017_1	Loan application	31413	35	0.41
14	bpic2017_2	Loan application	31413	35	0.12
15	bpic2017_3	Loan application	31413	35	0.47
16	Production	Manufacturing	220	9	0.53
17	sepsis_1	Hospital treatment	754	14	0.14
18	sepsis_2	Hospital treatment	782	13	0.14
19	sepsis_3	Hospital treatment	782	13	0.14
20	Traffic	Traffic fines	129615	4	0.46
21	hospital_1	Hospital finances	77525	6	0.1
22	hospital_2	Hospital finances	77525	6	0.05
23	insurance_1	Insurance	1065	12	0.16

Results: Learning Algorithms (Nemenyi test)

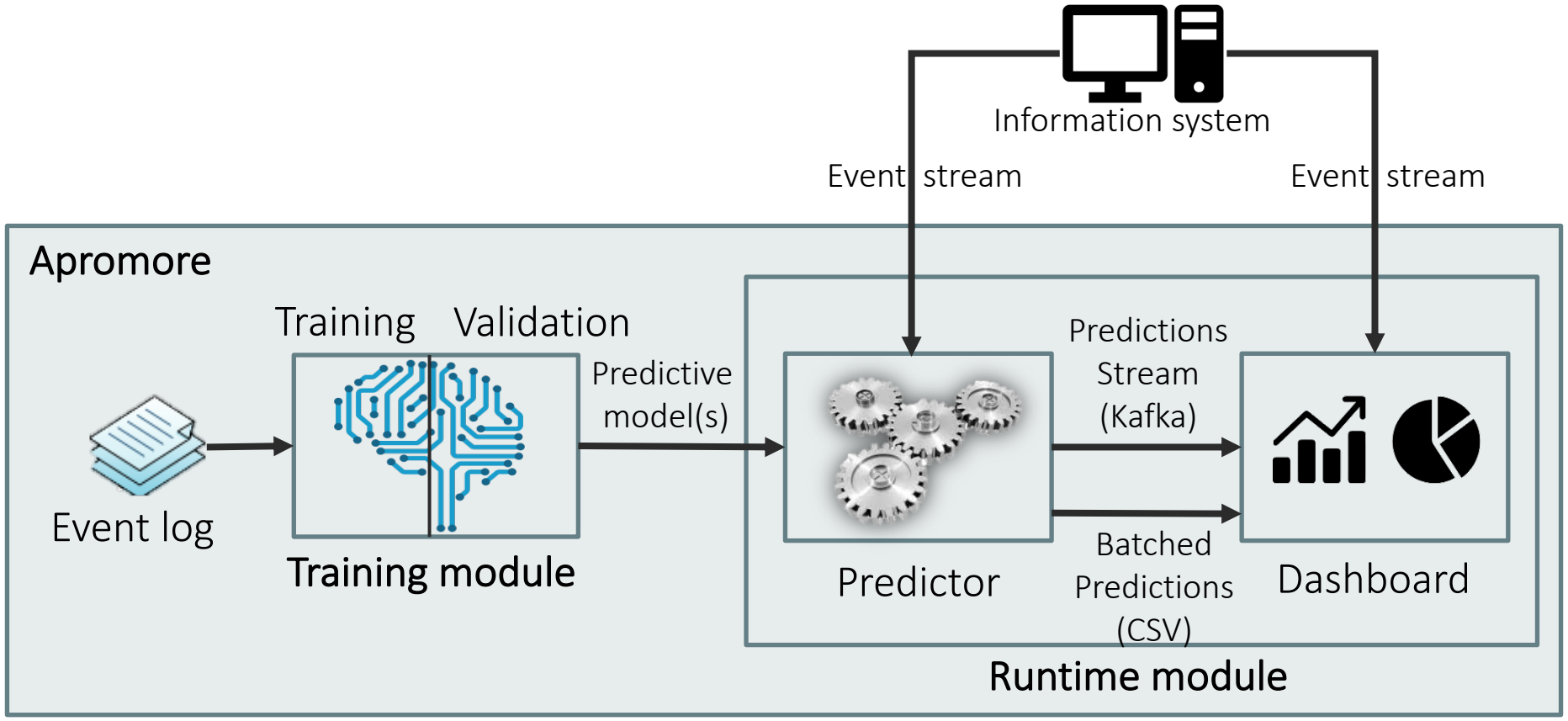


Results: Bucketing and Sequence Encoding

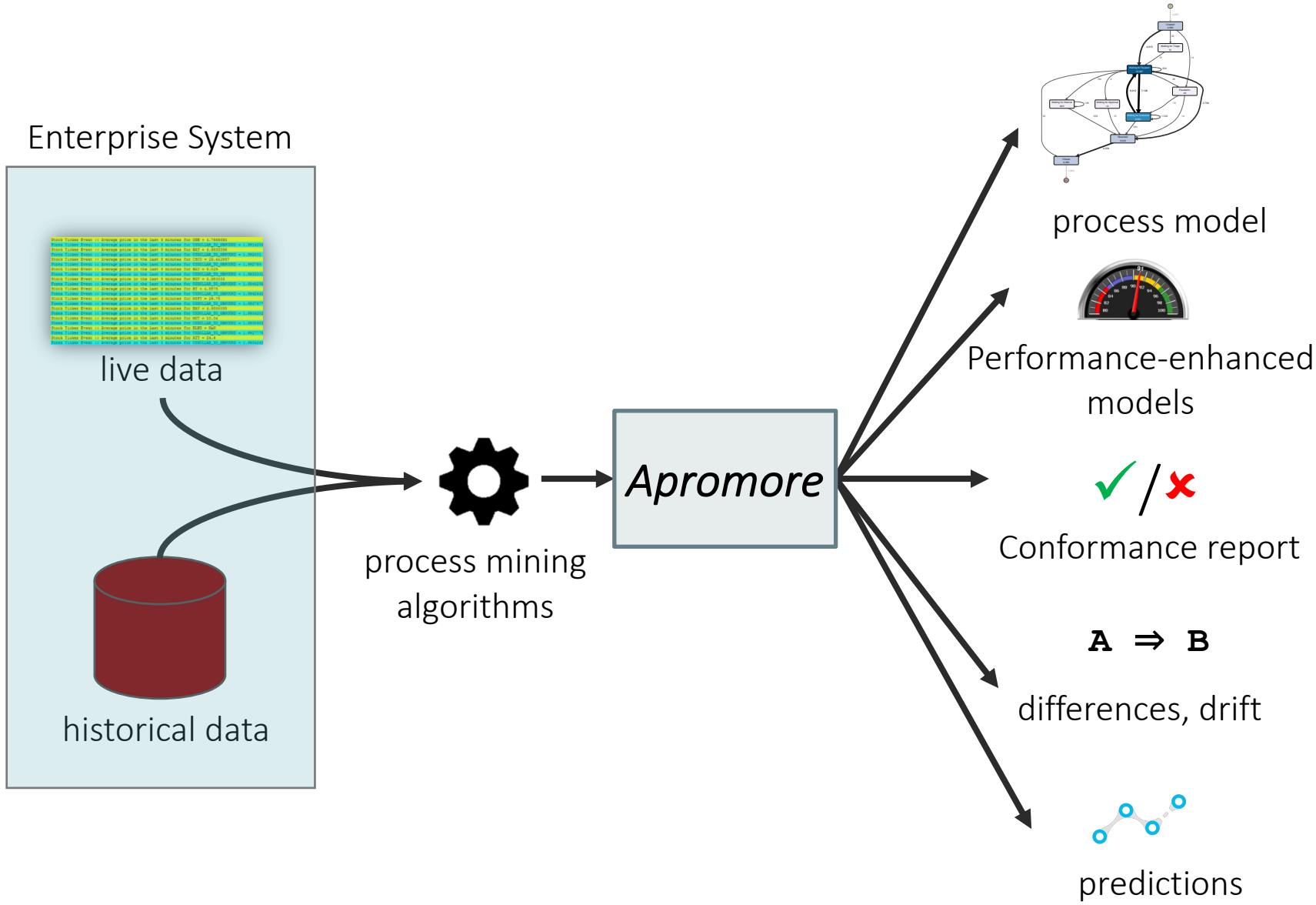


Predictive process monitoring in Apromore

- Predict **process outcome** (e.g. *"Is this loan offer going to be rejected?"*)
- Predict **process performance** (e.g. *"Will this claim take longer than 5 days to be handled?"*)
- Predict **future events** (e.g. *"What activity is likely to be executed next? And after that?"*)



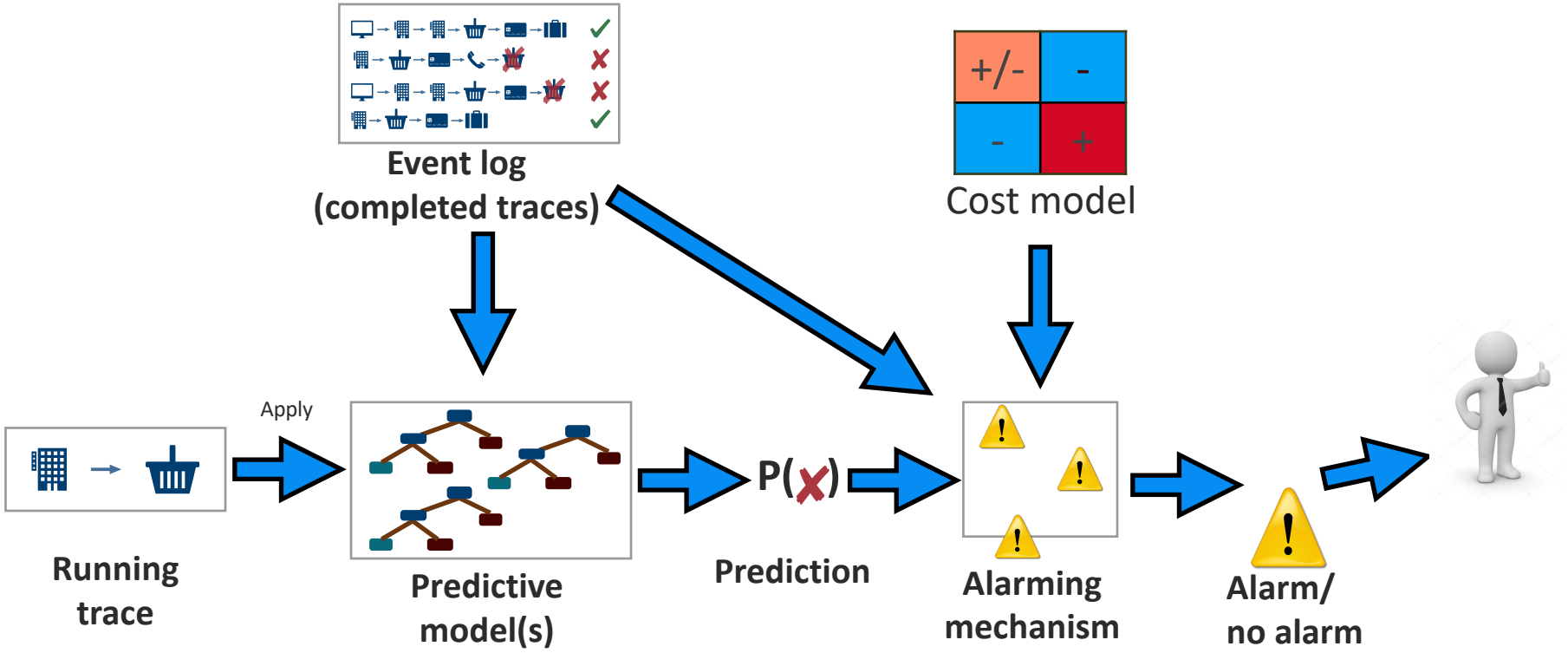
Recap: Process Mining in a Nutshell



Frontier topics in process mining

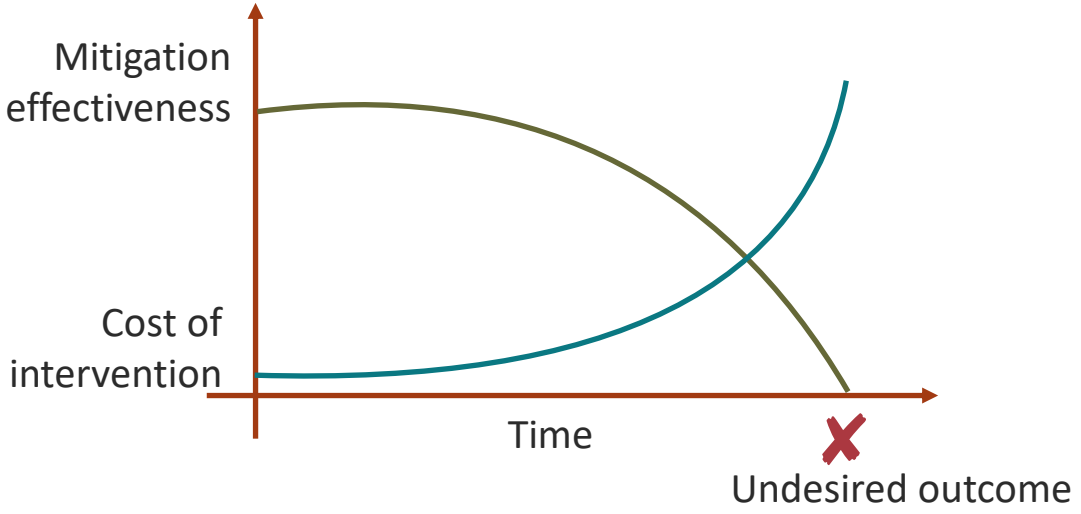
- **Explainable & Actionable Predictive Process Monitoring**
  - Extracting interpretable predictions
    - Helping users understand the root causes of predicted outcomes
  - Turning predictions into actions
    - **Prescriptive process monitoring**

Prescriptive process monitoring



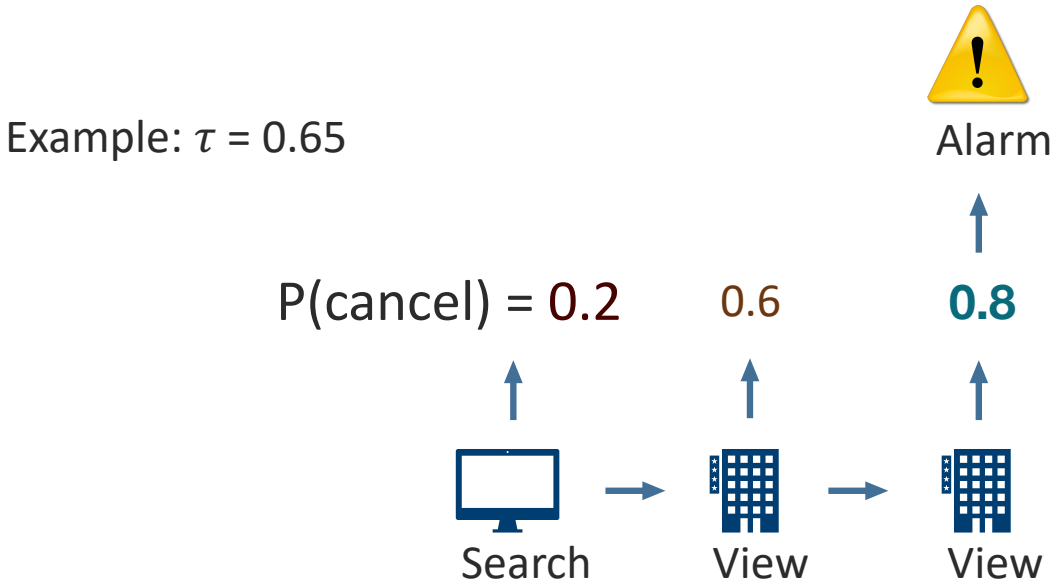
Cost model

	✗ Undesired outcome	✓ Desired outcome
⚠ Alarm raised	cost of intervention + (1 - mitigation effectiveness) * cost of undesired outcome	cost of intervention + cost of compensation
Alarm not raised	cost of undesired outcome	no costs

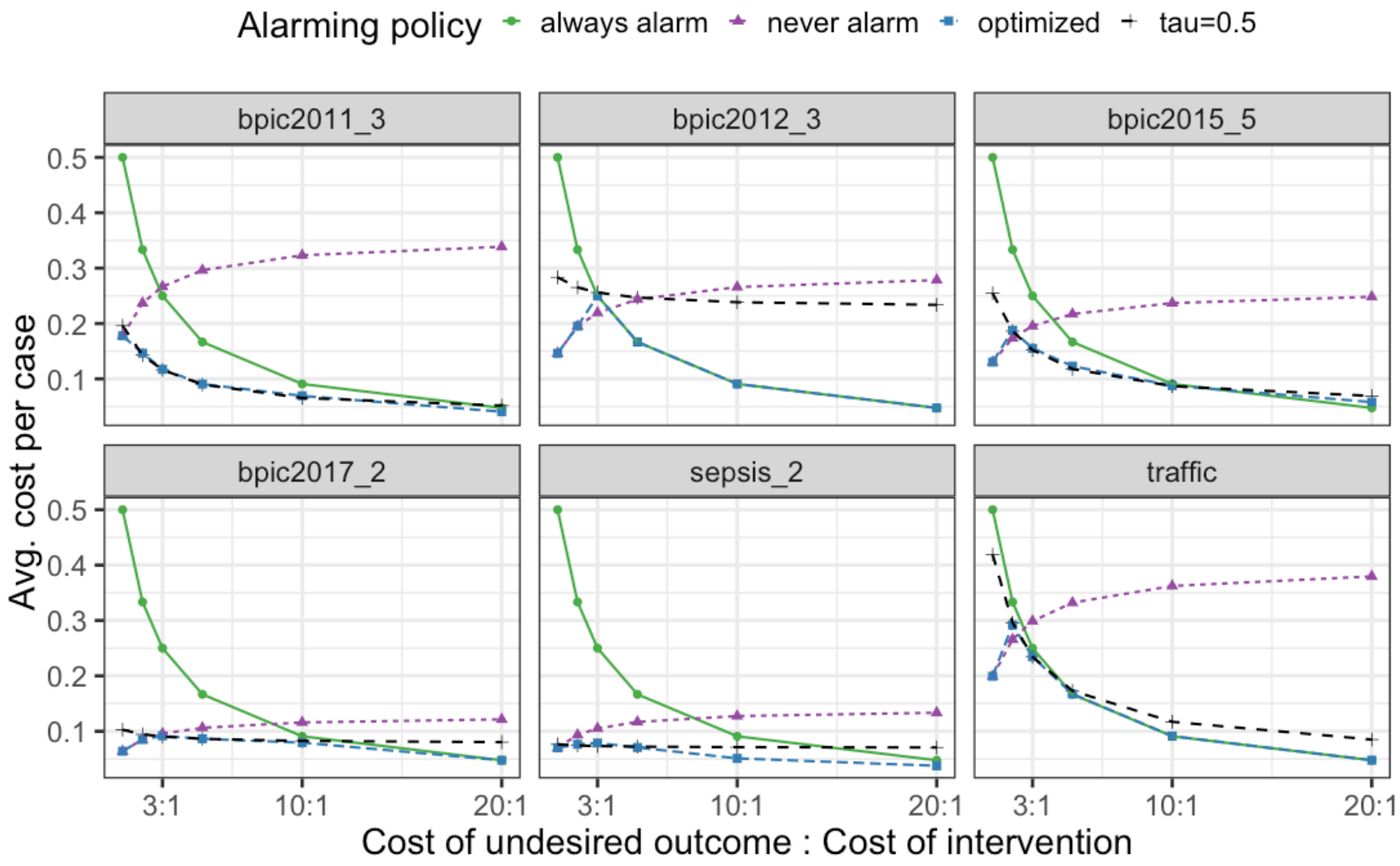


Alarming mechanism

- Raise an alarm if P(undesired outcome) > τ
- Optimal τ is found via empirical thresholding



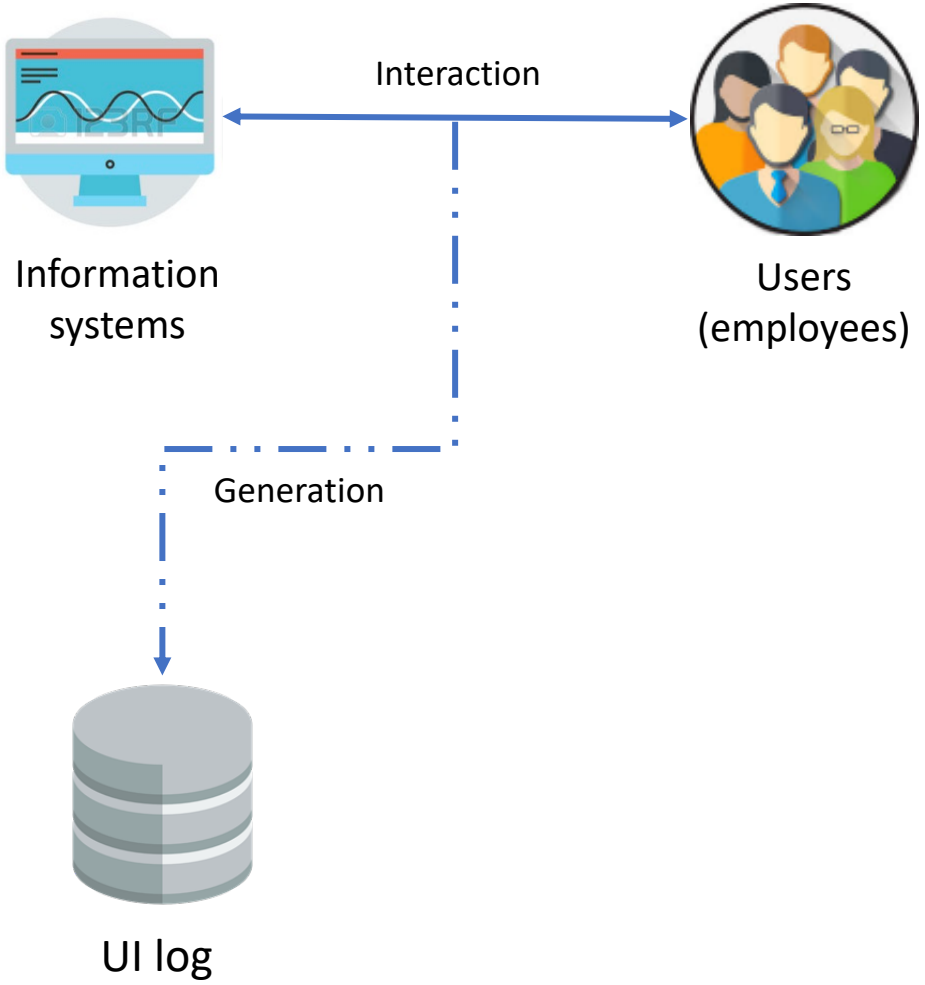
Results


Alarming policy
always alarm
never alarm
optimized
tau=0.5
bpic2011_3
bpic2012_3
bpic2015_5
bpic2017_2
sepsis_2
traffic
Avg. cost per case
0.5
0.4
0.3
0.2
0.1
3:1
10:1
20:1
Cost of undesired outcome : Cost of intervention

Frontier topics in process mining

- **Explainable & Actionable Predictive Process Monitoring**
  - Extracting interpretable predictions
    - Helping users understand the root causes of predicted outcomes
  - Turning predictions into actions
    - **Prescriptive process monitoring**
- **Robotic Process Mining**
  - Discovering executable routine specifications (e.g. RPA scripts) from UI logs

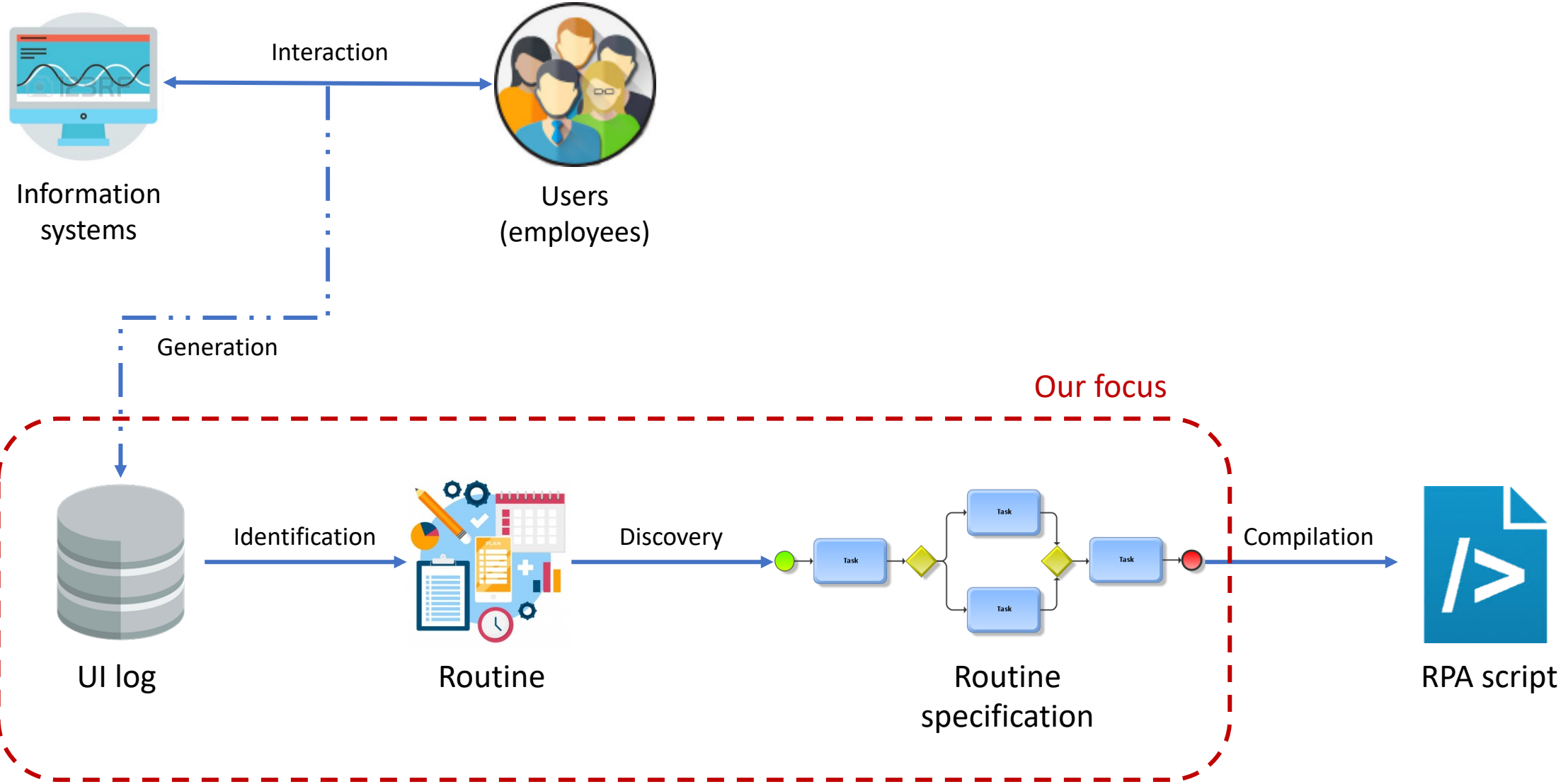
Robotic Process Mining



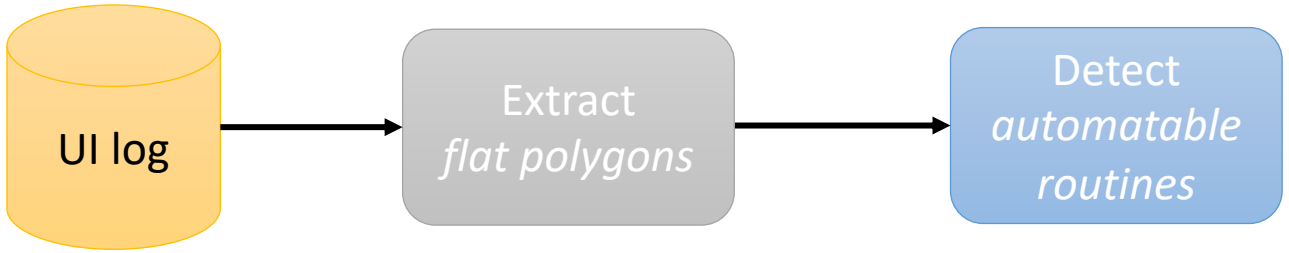
UI log

	Action Type	Action Parameters			
		Param-1	**Param-2**	**Param-3**	**Param-4**
1	Click button	Target:Web	Label: STUDENTS		
2	Fill the text field	Target:Web	Label: ID Student	Value: 010234	
3	Press key	Target:Web	Label: ENTER		
4	Click button in row	Target:Web	Label: Update	ID Row: 010234	
5	Fill the text field	Target:Web	Label: Address	Value: 19 Parkville St, Burnley VIC 3121	
6	Fill the text field	Target:Web	Label: Country	Value: Australia	
7	Open file	Target:Excel	Name: 010234	Path: C:/Students/Australia/	Extension: .xls
8	Copy (Ctrl+C)	Target:Web	From: Address	Value: 19 Parkville St	
9	Paste (Ctrl+V)	Target:Excel	Row: 5	Column: A	Value: 19 Parkville St
10	Save file (Ctrl+S)	Target:Excel			
11	Click button	Target:Web	Label: Confirm Backup		

Robotic Process Mining



Robotic Process Mining: Initial Approach

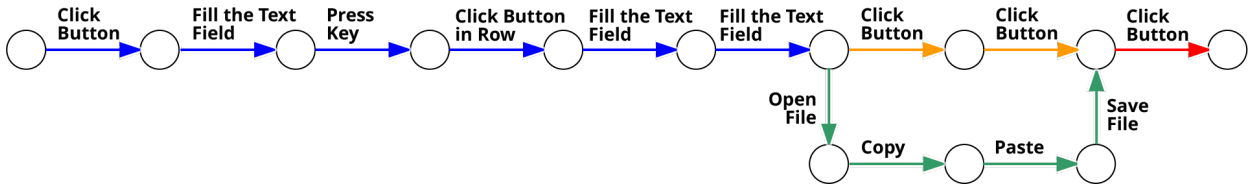


Detect automatable routines:

1. Detect automatable actions
2. Return only those flat polygons made of automatable actions

An action is *automatable* if all its arguments are *constant* or *functions* of arguments of previously-executed actions

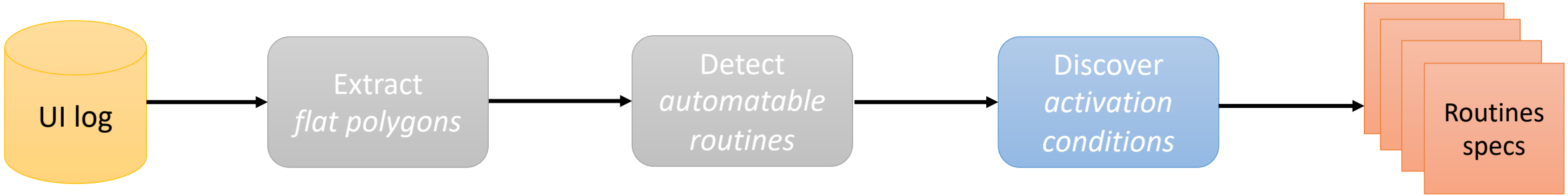
Robotic Process Mining: Initial Approach



Foofah – Discovering Data Transformations by Example

	Action Type	Action Parameters			
		Param-1	**Param-2**	**Param-3**	**Param-4**
1	Click button	Target:Web	Label: STUDENTS		
2	Fill the text field	Target:Web	Label: ID Student	Value: 010234	
3	Press key	Target:Web	Label: ENTER		
4	Click button in row	Target:Web	Label: Update	ID Row: 010234	
5	Fill the text field	Target:Web	Label: Address	Value: 19 Parkville St Burnley VIC 3121	
6	Fill the text field	Target:Web	Label: Country	Value: Australia	
7	Open file	Target:Excel	Name: 010234	Path: C:/Students/Australia/	Extension: .xls
8	Copy (Ctrl+C)	Target:Web	From: Address	Value: 19 Parkville St	
9	Paste (Ctrl+V)	Target:Excel	Row: 5	Column: A	Value: 19 Parkville St
10	Save file (Ctrl+S)	Target:Excel			
11	Click button	Target:Web	Label: Confirm Backup		

Robotic Process Mining: Initial Approach



<u>Discover activation conditions:</u>

For each automatable routine, discover its *activation condition,* containing:

1. *Triggering action*, which must be successfully executed before the routine
2. *Boolean condition*, which must be valid at the completion of the triggering action

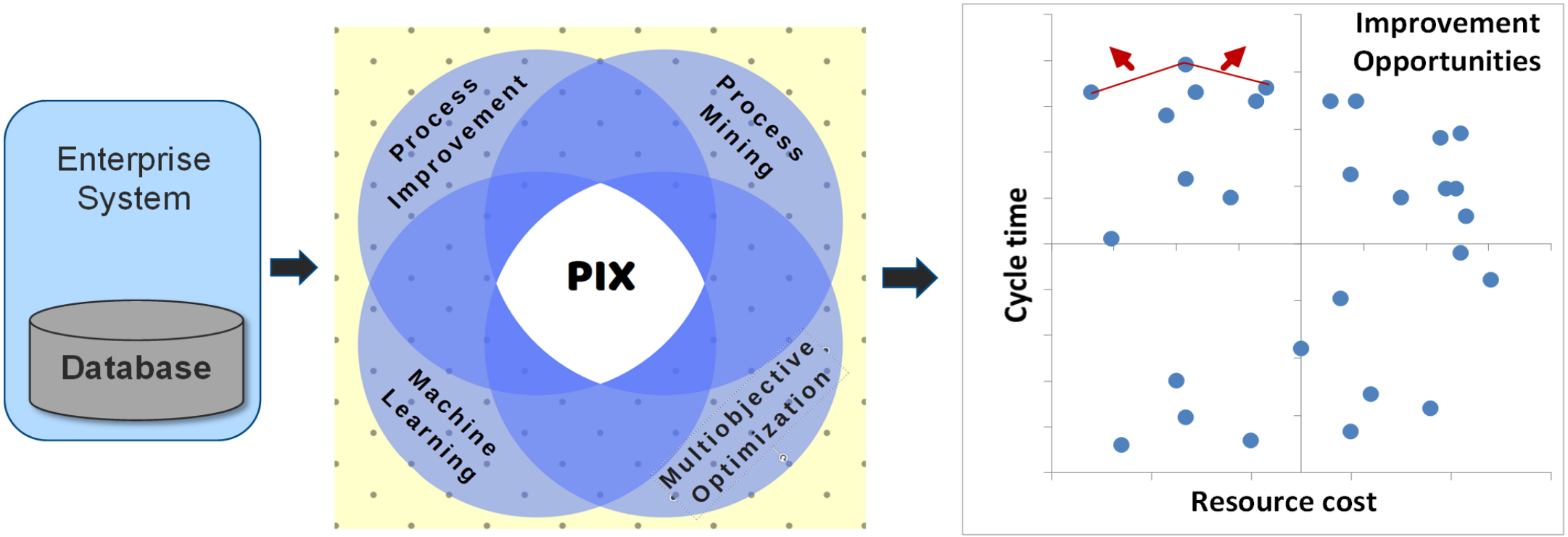
Robotic Process Mining: Initial Approach

	Action Type	Action Parameters			
		Param-1	**Param-2**	**Param-3**	**Param-4**
1	Click button	Target:Web	Label: STUDENTS		
2	Fill the text field	Target:Web	Label: ID Student	Value: 010234	
3	Press key	Target:Web	Label: ENTER		
4	Click button in row	Target:Web	Label: Update	ID Row: 010234	
5	Fill the text field	Target:Web	Label: Address	Value: 19 Parkville St, Burnley VIC 3121	
6	Fill the text field	Target:Web	Label: Country	Value: Australia	
7	Open file	Target:Excel	Name: 010234	Path: C:/Students/Australia/	Extension: .xls
8	Copy (Ctrl+C)	Target:Web	From: Address	Value: 19 Parkville St	
9	Paste (Ctrl+V)	Target:Excel	Row: 5	Column: A	Value: 19 Parkville St
10	Save file (Ctrl+S)	Target:Excel			
11	Click button	Target:Web	Label: Confirm Backup		

Frontier topics in process mining

- **Explainable & Actionable Predictive Process Monitoring**
  - Extracting interpretable predictions
    - Helping users understand the root causes of predicted outcomes
  - Turning predictions into actions
    - **Prescriptive process monitoring**
- **Robotic Process Mining**
  - Discovering executable routine specifications (e.g. RPA scripts) from UI logs
- **Automated Process Discovery**
  - Given an event log L, discover opportunities to improve the process w.r.t. one or more performance measures

The Process Improvement Explorer (PIX)



WE ARE HIRING!

2 x PhD positions & 2 x Postdoc Postions

https://sep.cs.ut.ee/Main/PIX

erc