

Mining Configurable Process Models from Collections of Event Logs

Joos Buijs

Boudewijn van Dongen

Wil van der Aalst

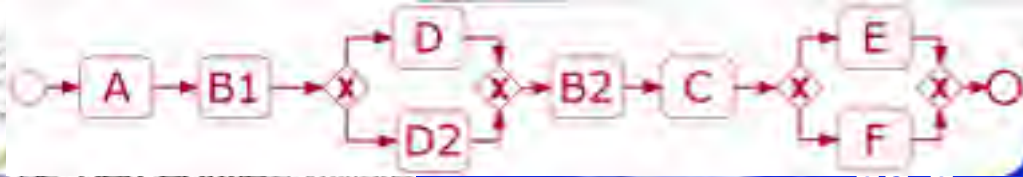
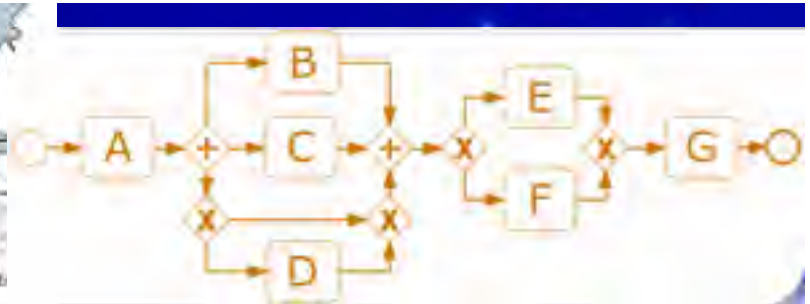
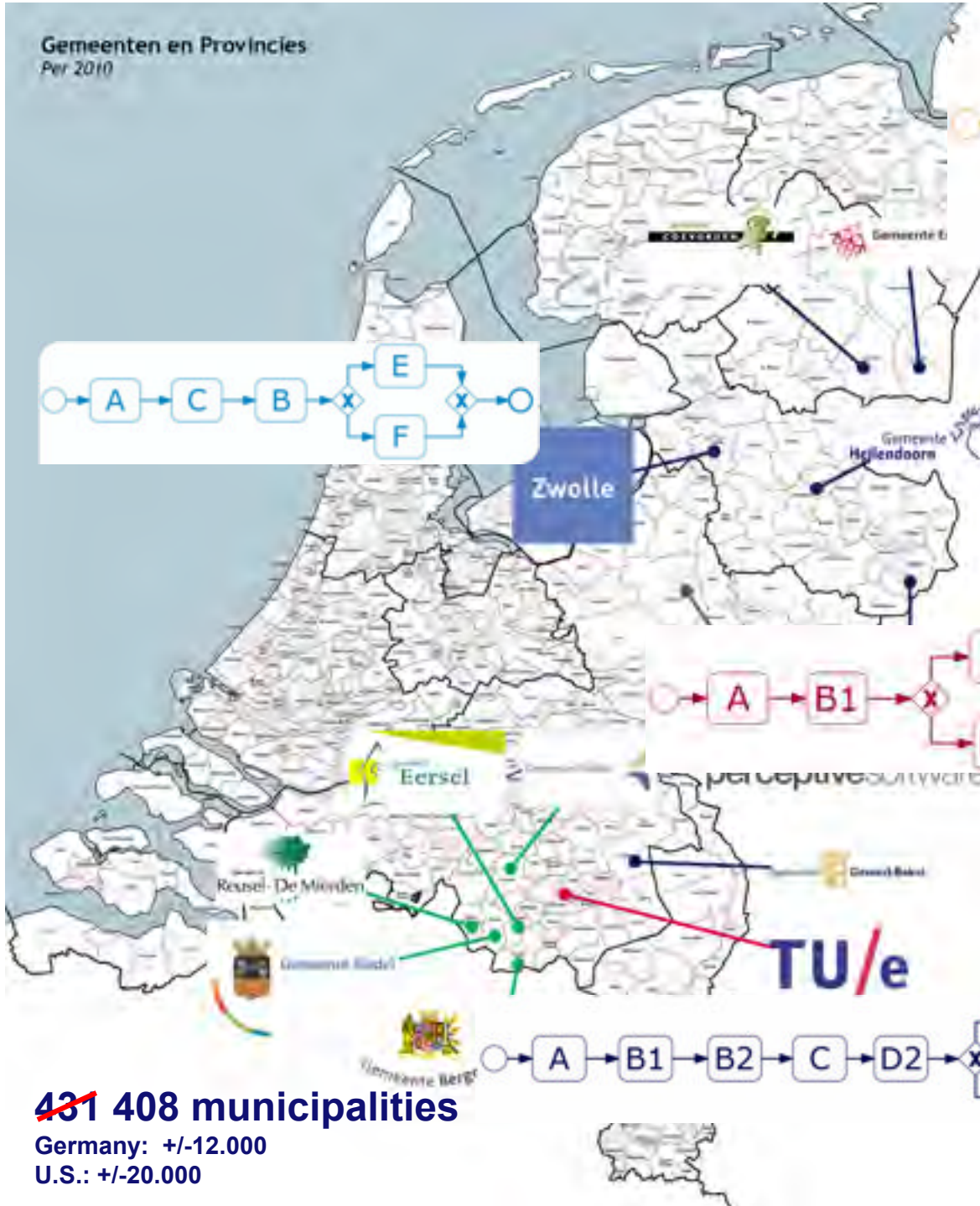


TU/e

Technische Universiteit
Eindhoven
University of Technology

Where innovation starts

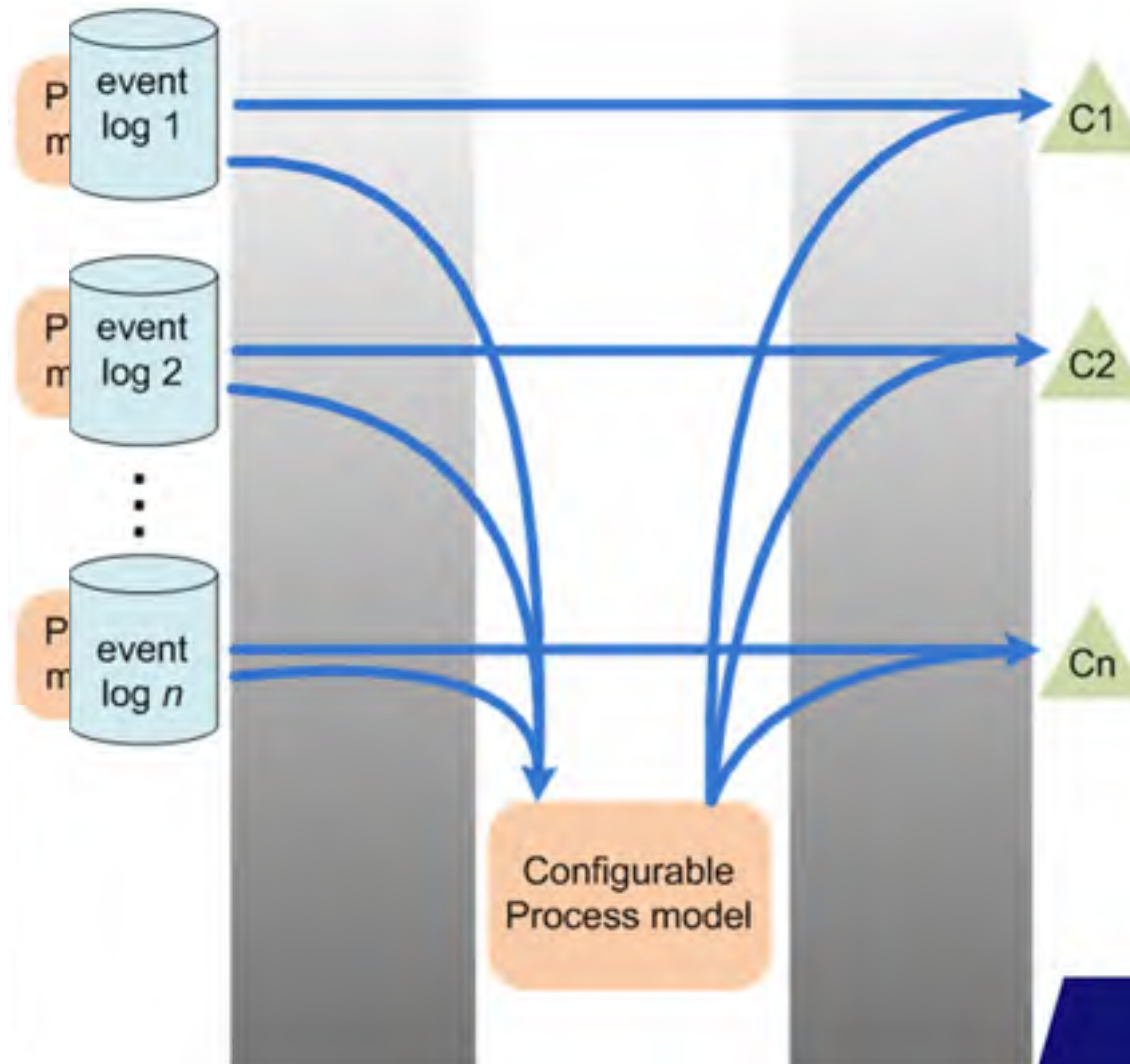
Gemeenten en Provincies
Per 2010



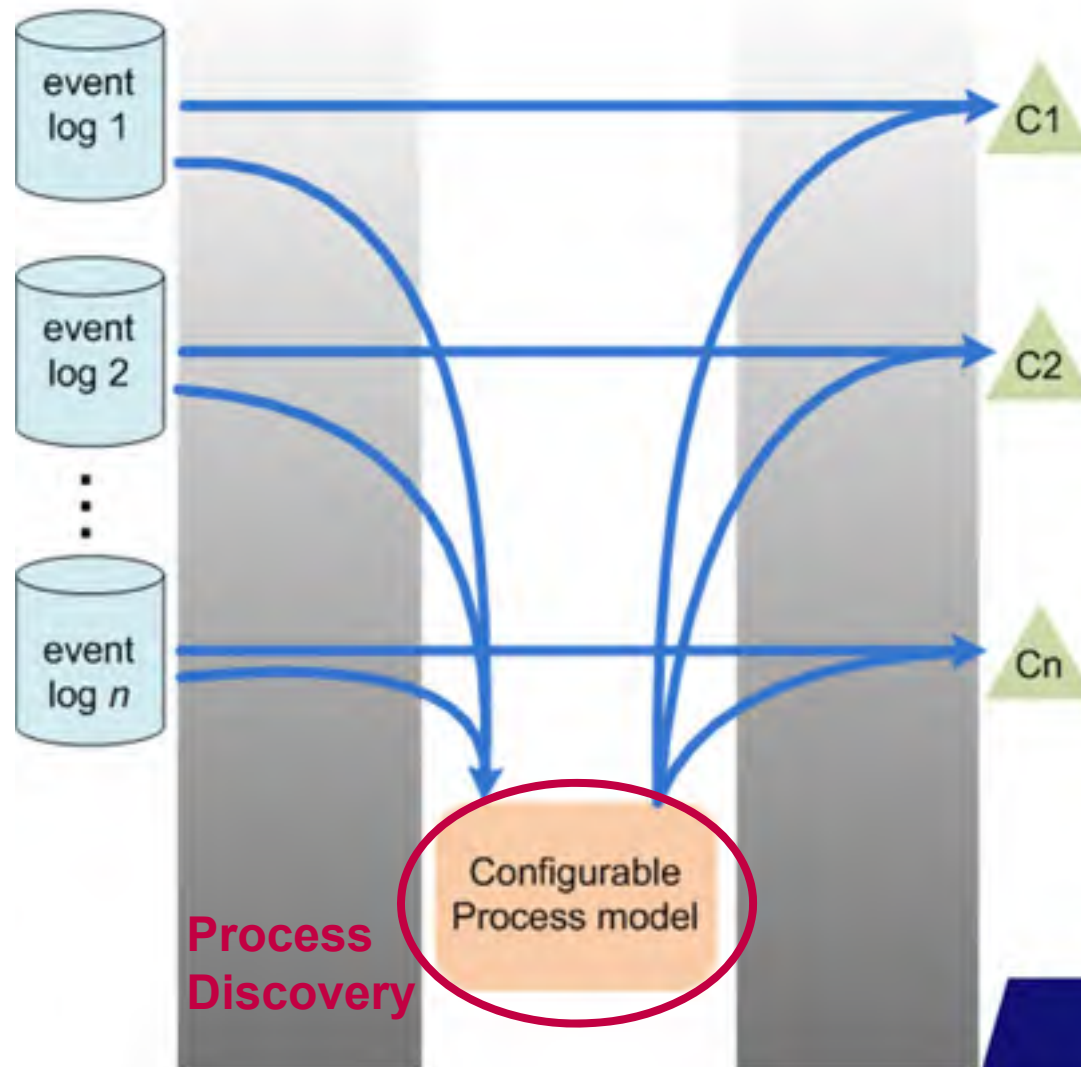
as a
Service
(SaaS)

~~431~~ 408 municipalities
Germany: +/-12.000
U.S.: +/-20.000

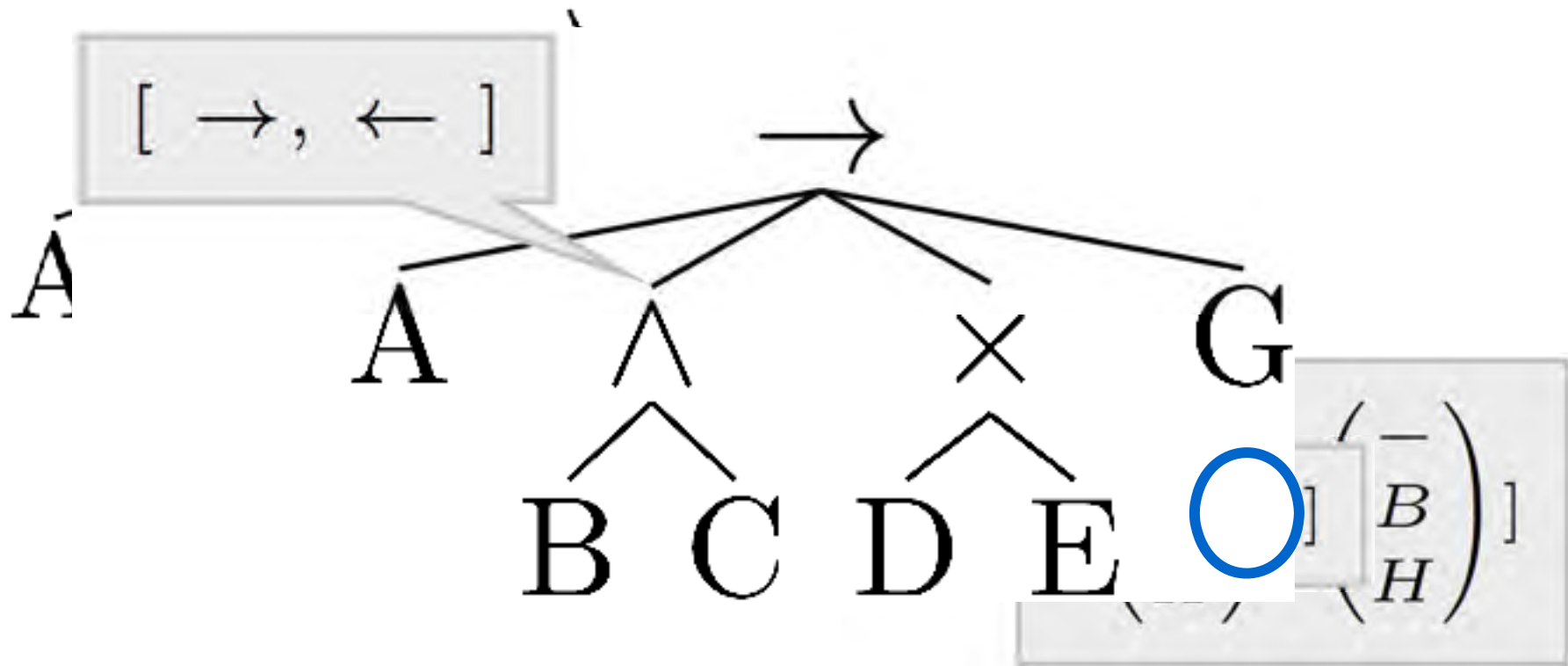
Creating a Configurable Process Model



Overview

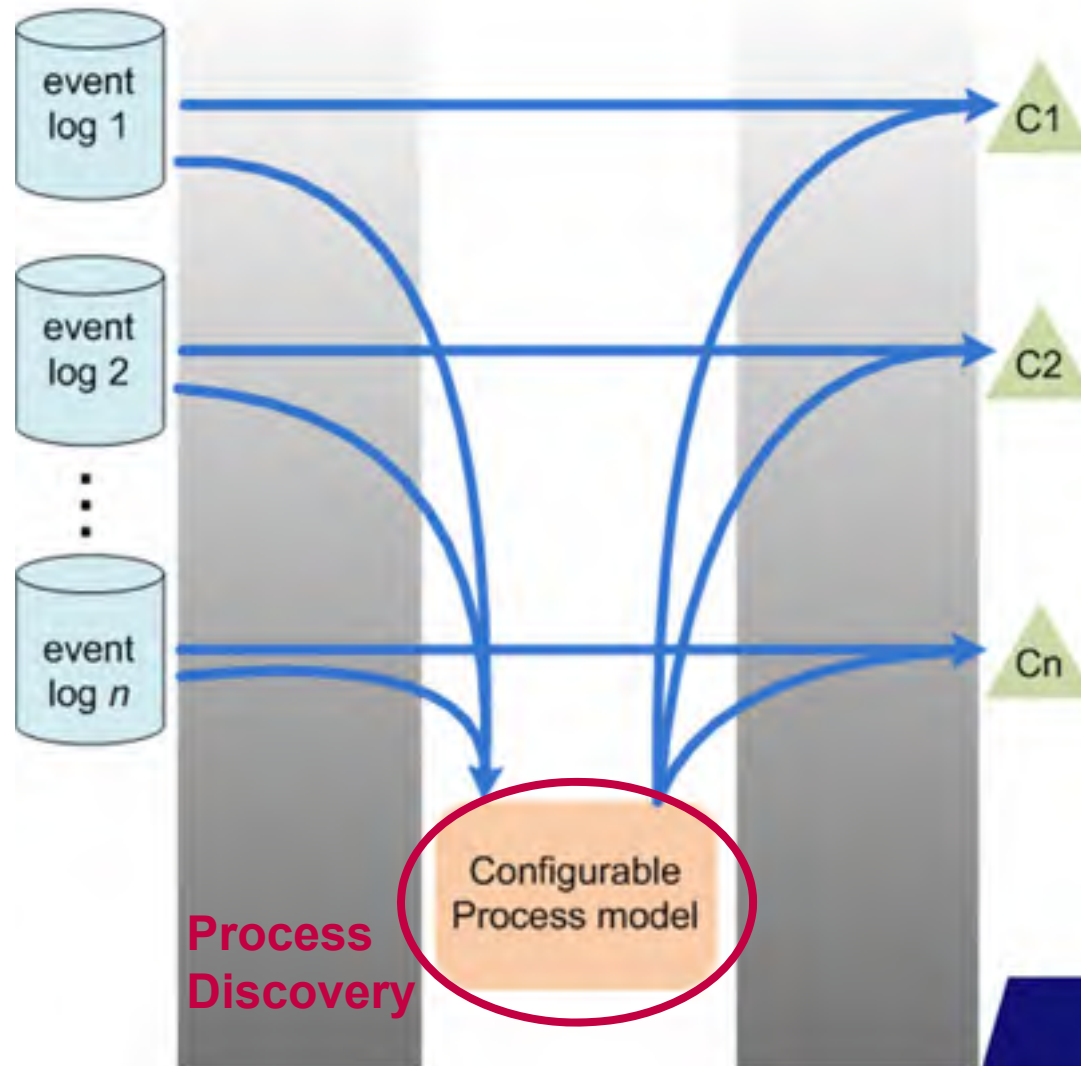


Configurable Process Trees



- Configuration points with options that can be set
- Hiding
- Blocking
- Operator Downgrade

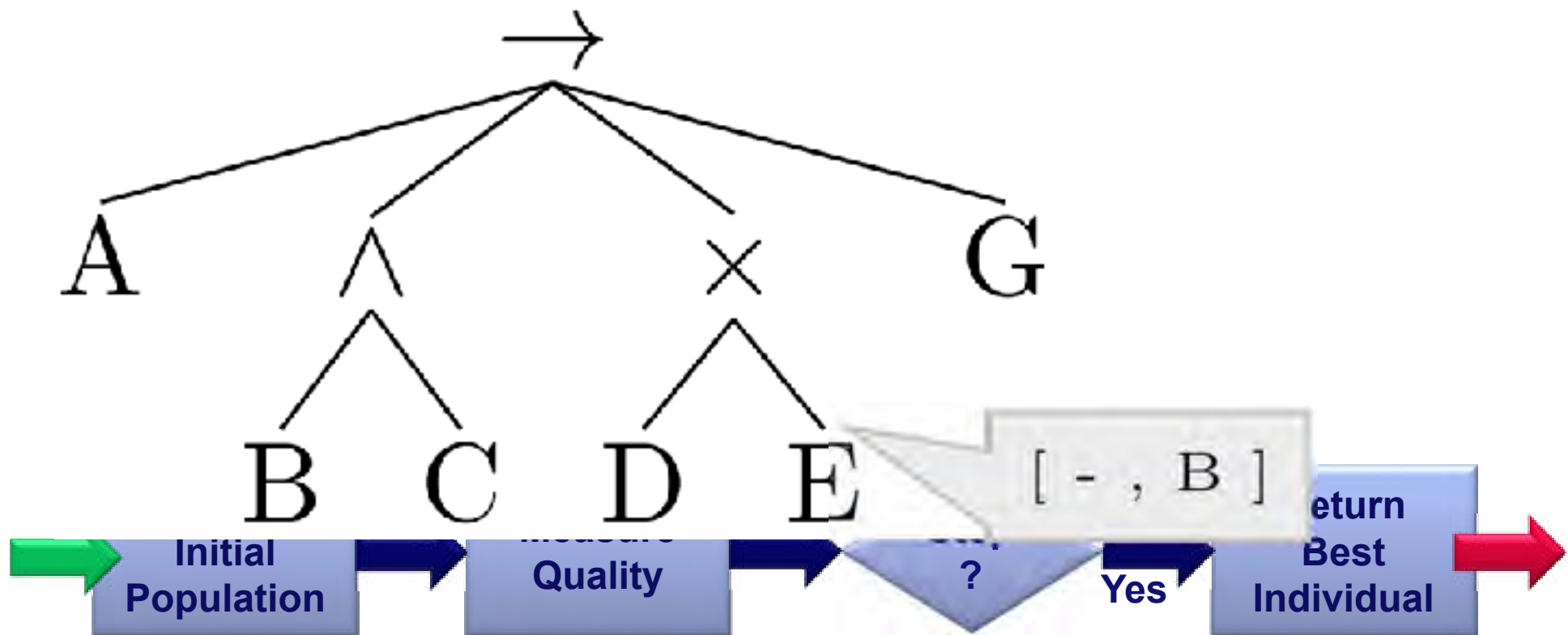
Overview



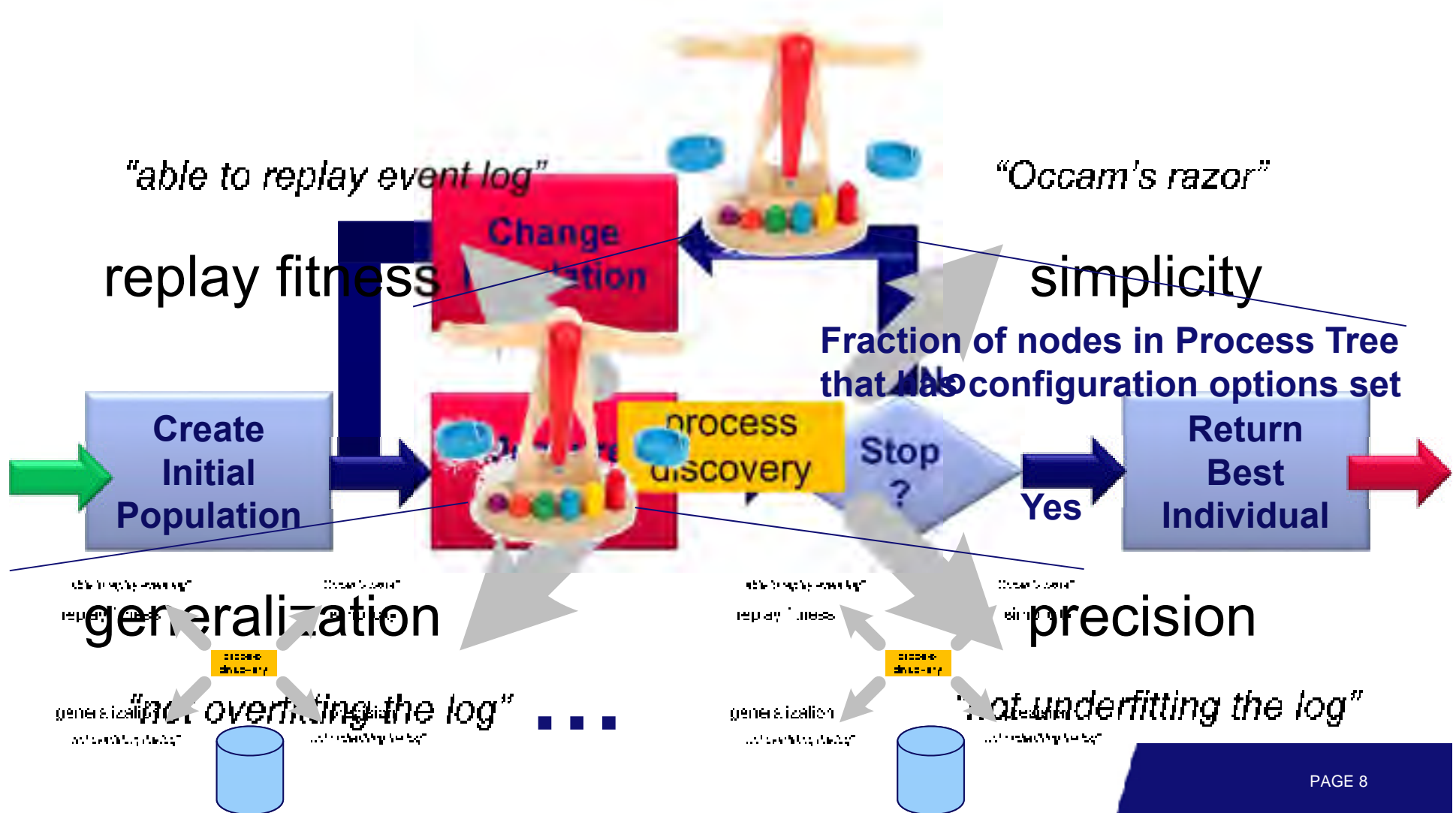
The Evolutionary Tree Miner Algorithm



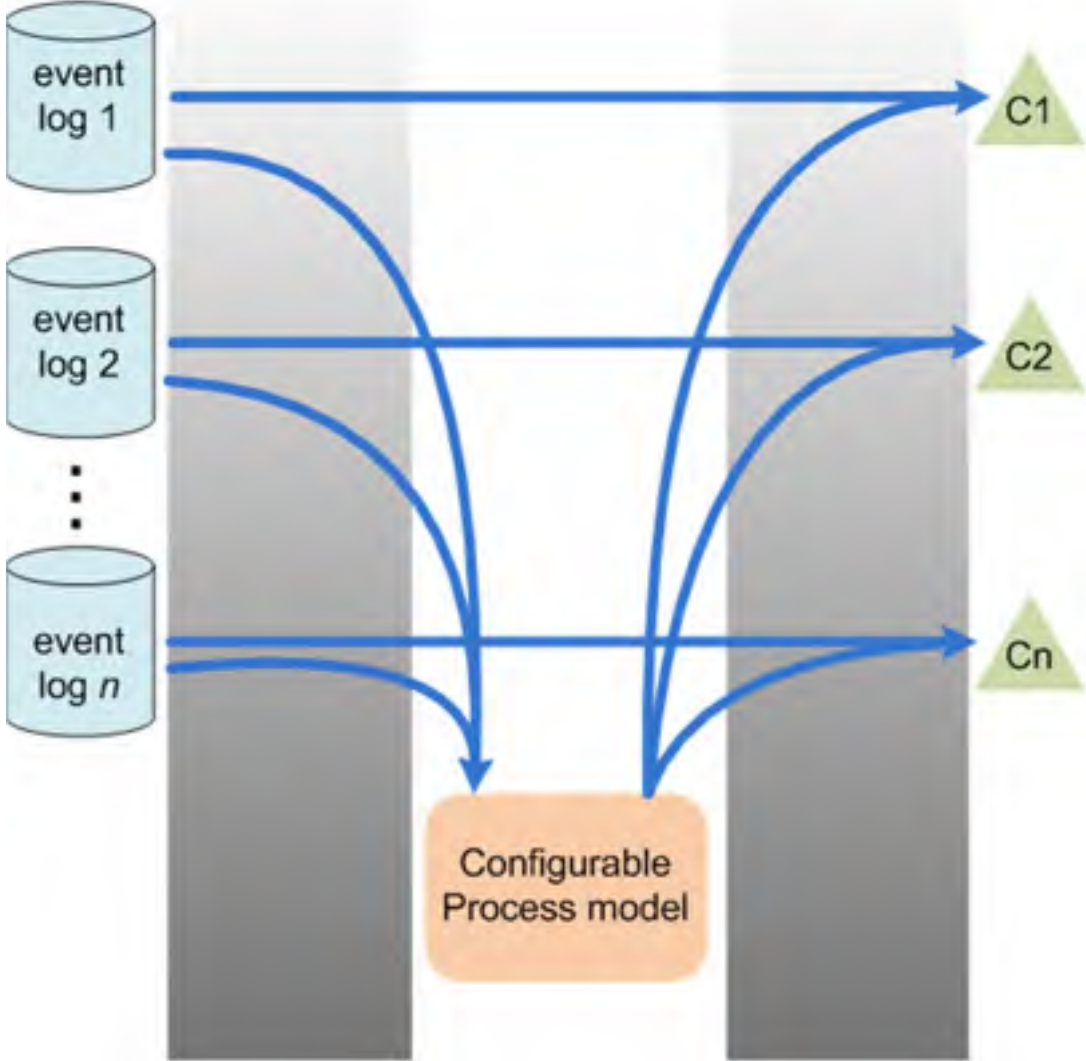
The ETM Algorithm – Population Change



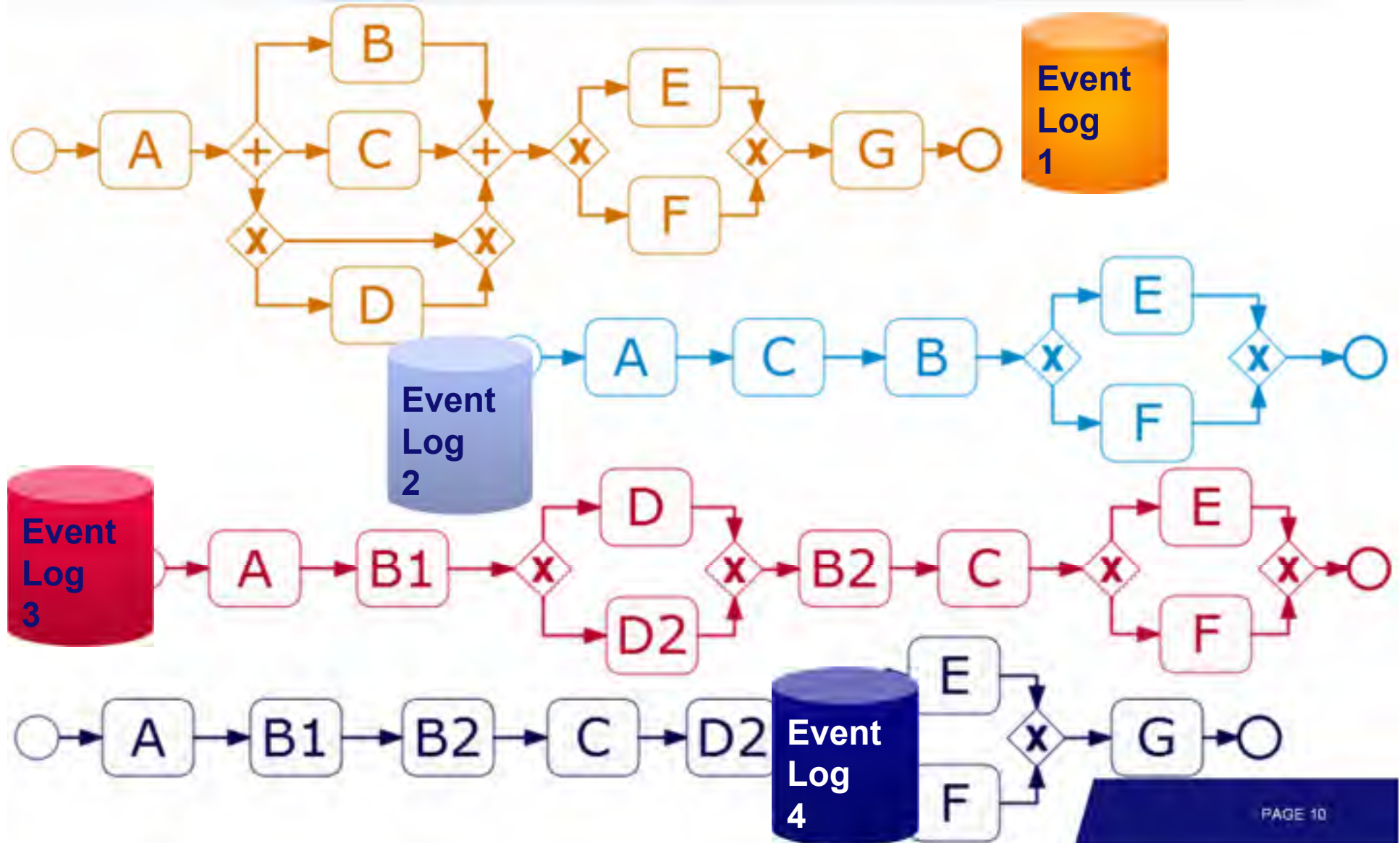
The ETM Algorithm – Measuring Quality



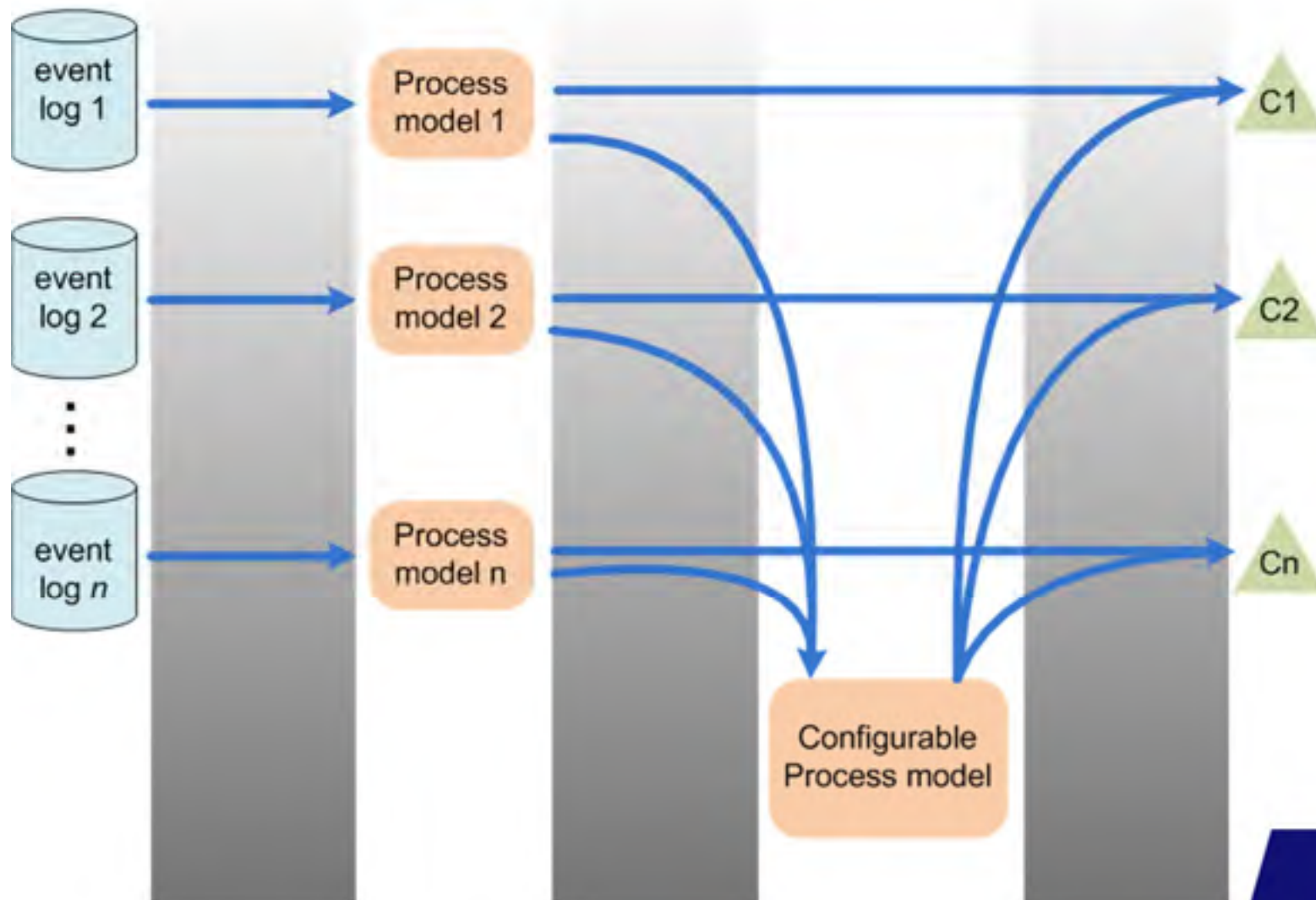
Overview



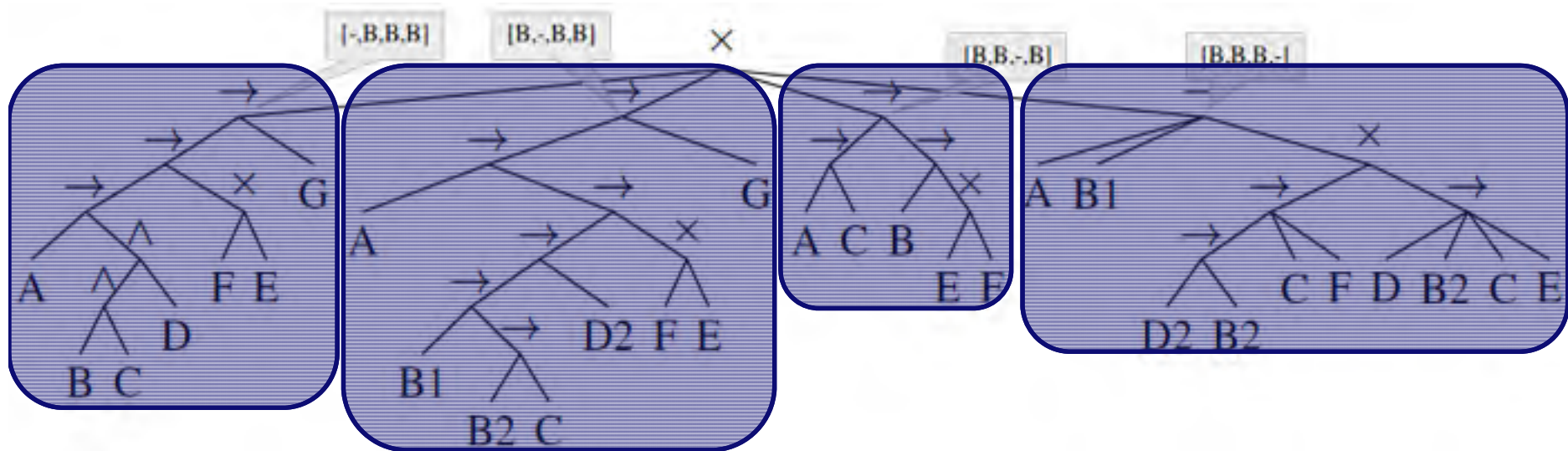
Running Example: 4 Variants



Approach 1 (Gottschalk et. al.)

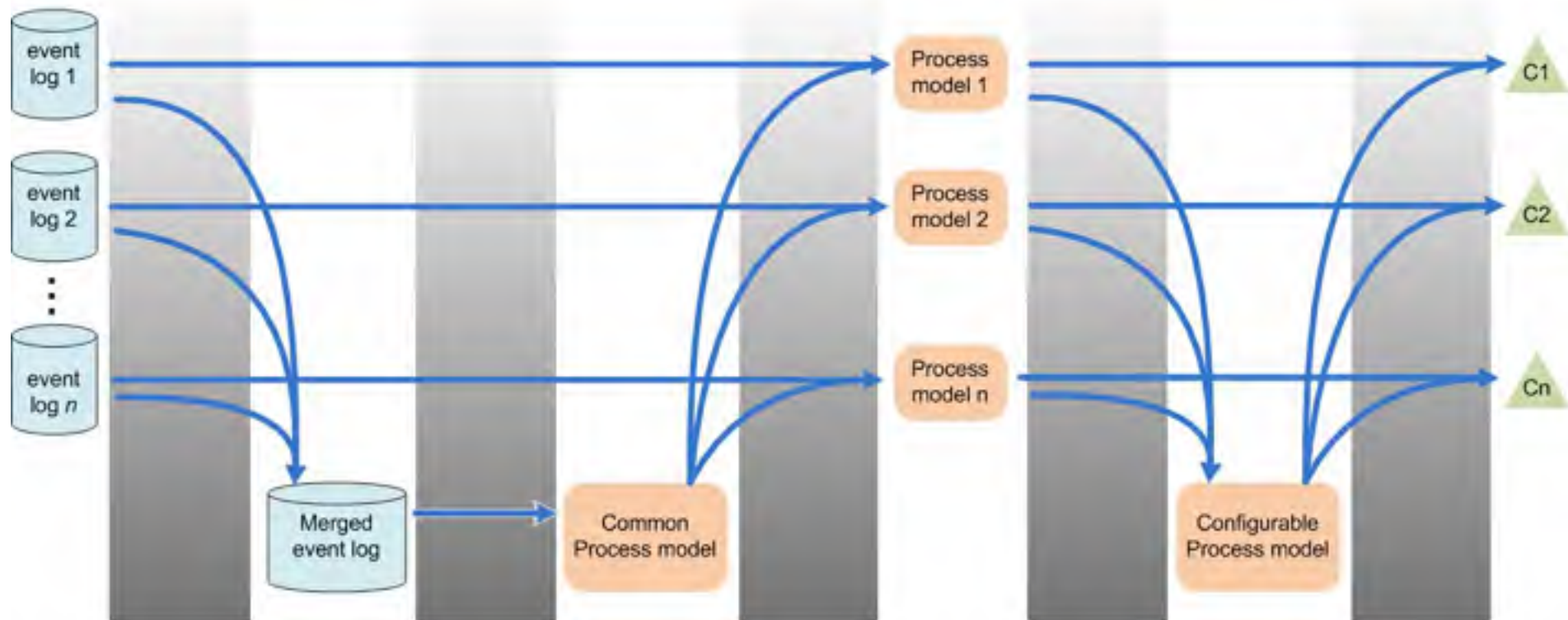


Result Approach 1

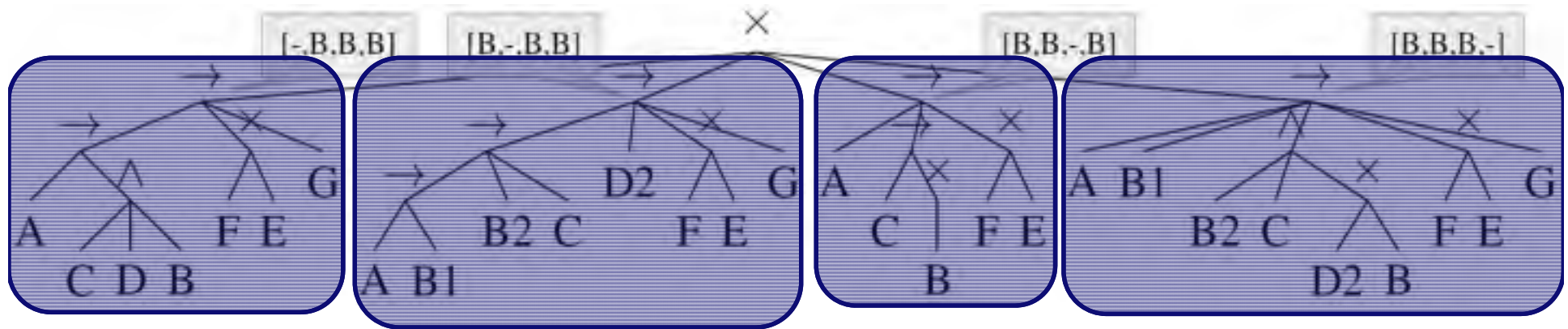


Common parts should be shared but are hard to detect

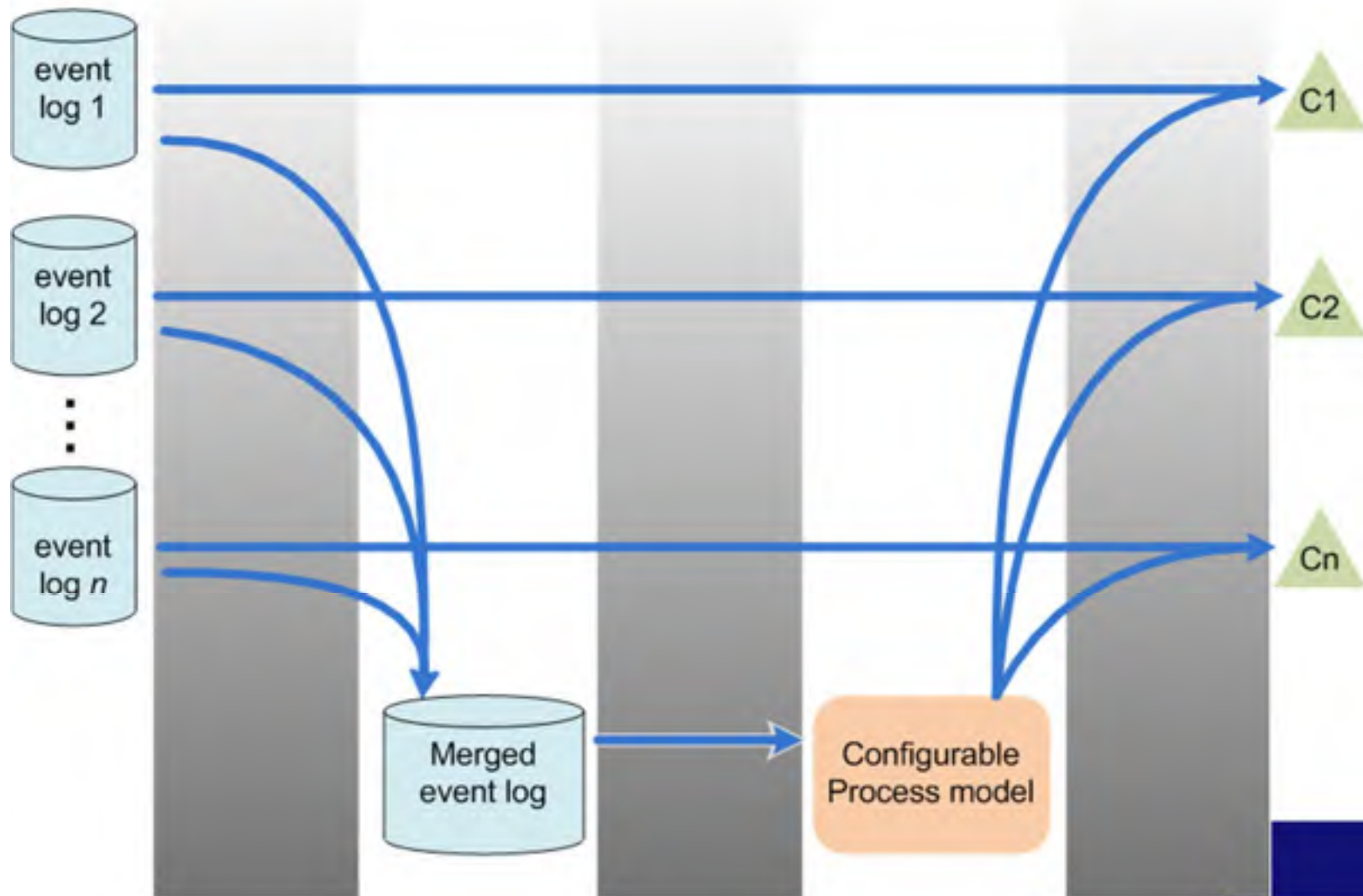
Approach 2 – More Similar Input Models



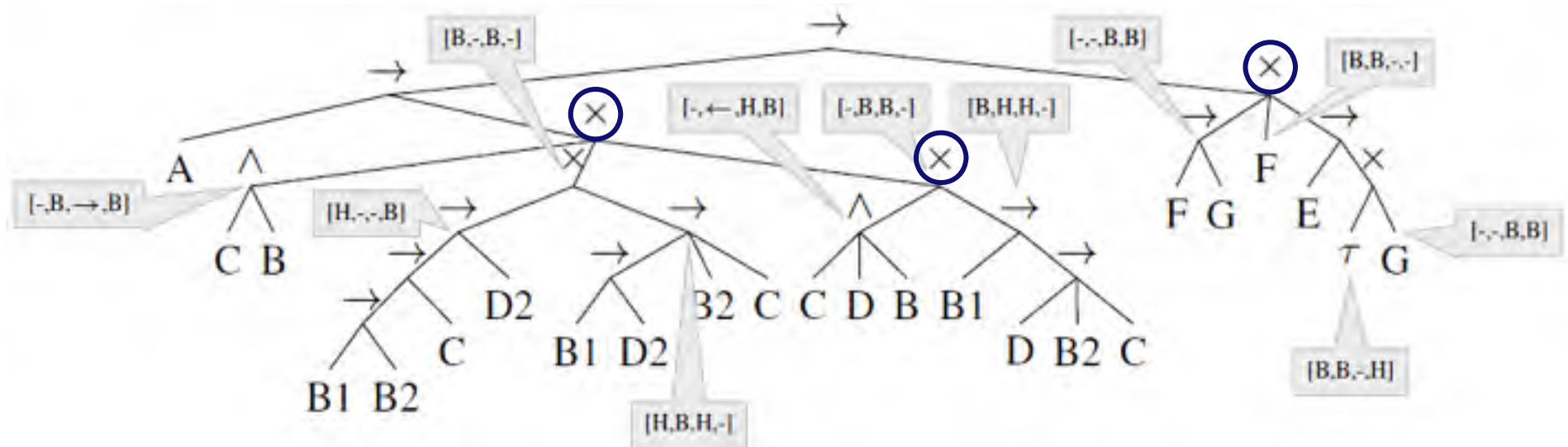
Result Approach 2



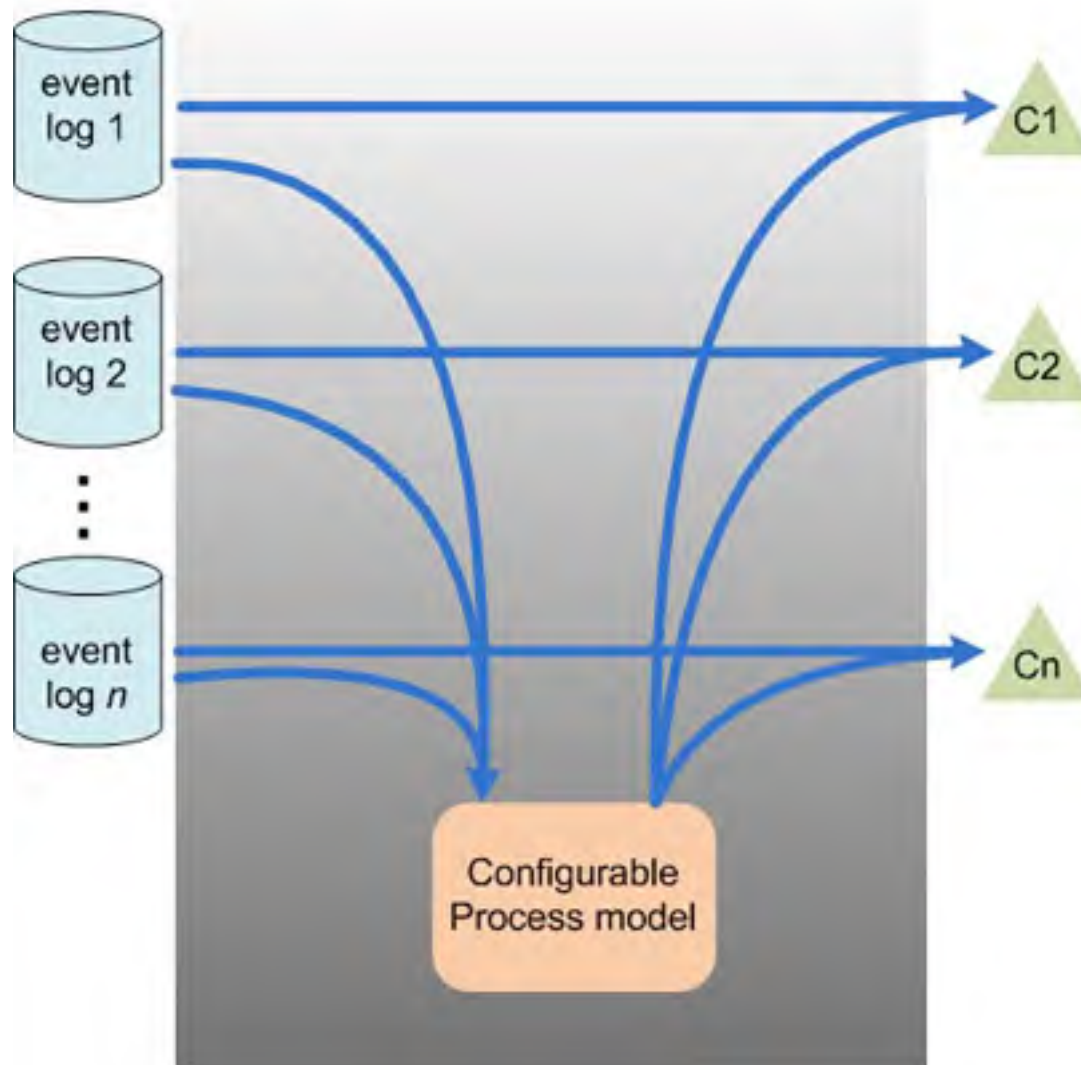
Approach 3 (Gottschalk et. al.)



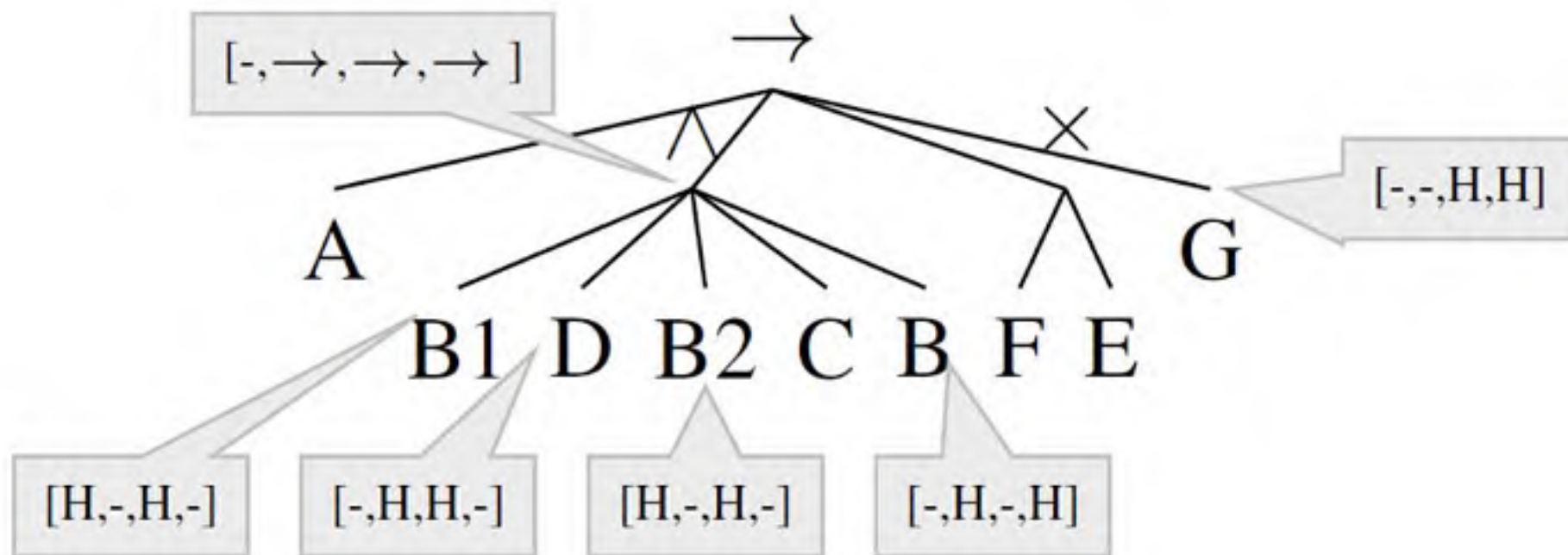
Approach 3: Result



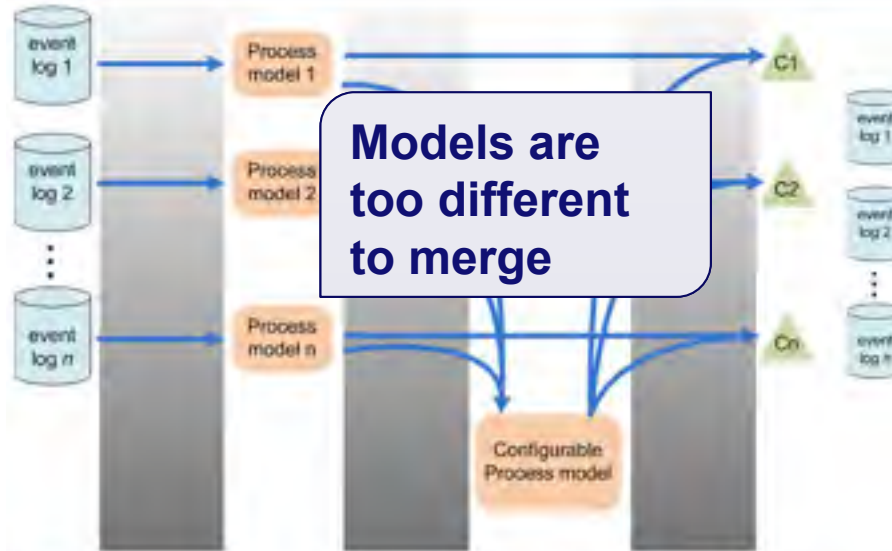
Approach 4



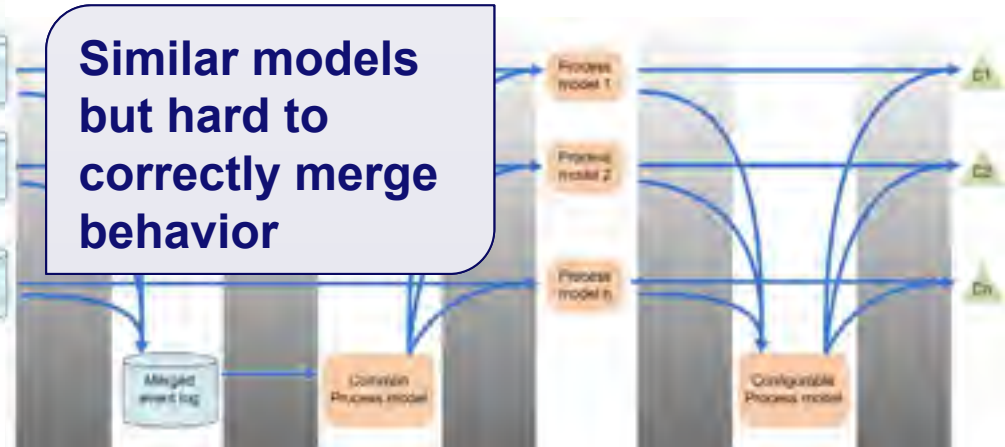
Approach 4: Result



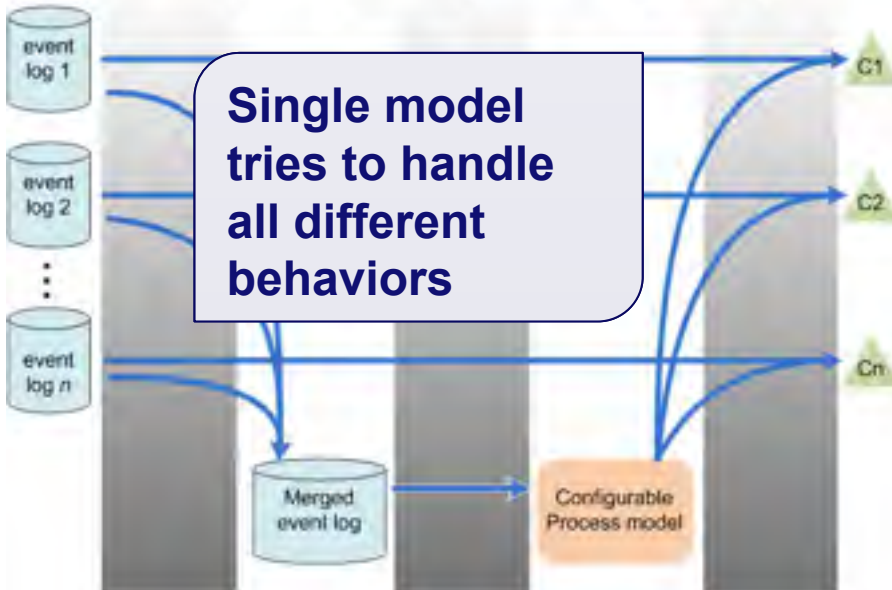
Conclusion



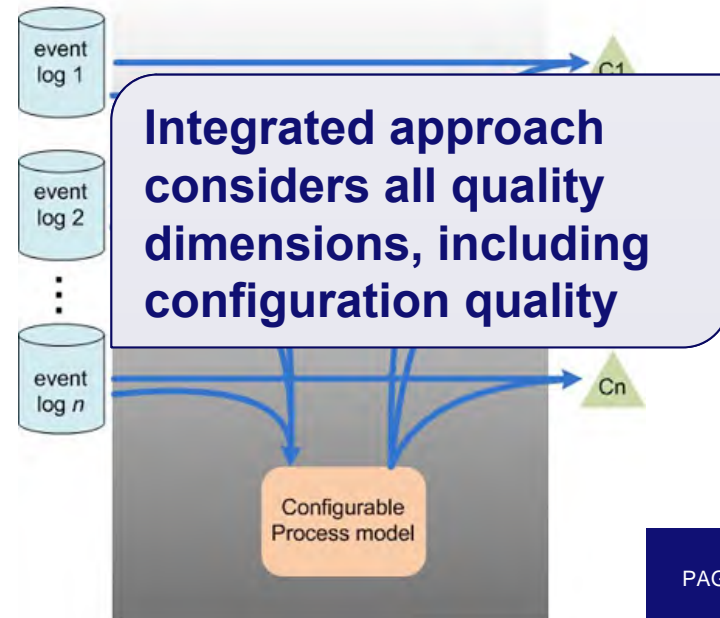
Models are too different to merge



Similar models but hard to correctly merge behavior



Single model tries to handle all different behaviors



Integrated approach considers all quality dimensions, including configuration quality

Take Home Points

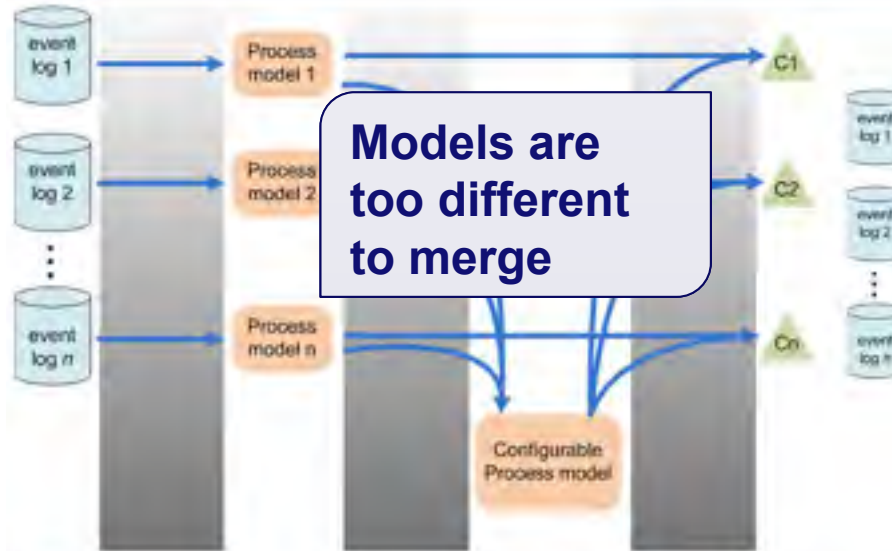
- Merging process models on their behavior is difficult
- Process Discovery is all about abstracting behavior
 - But the input event logs should be kept separate



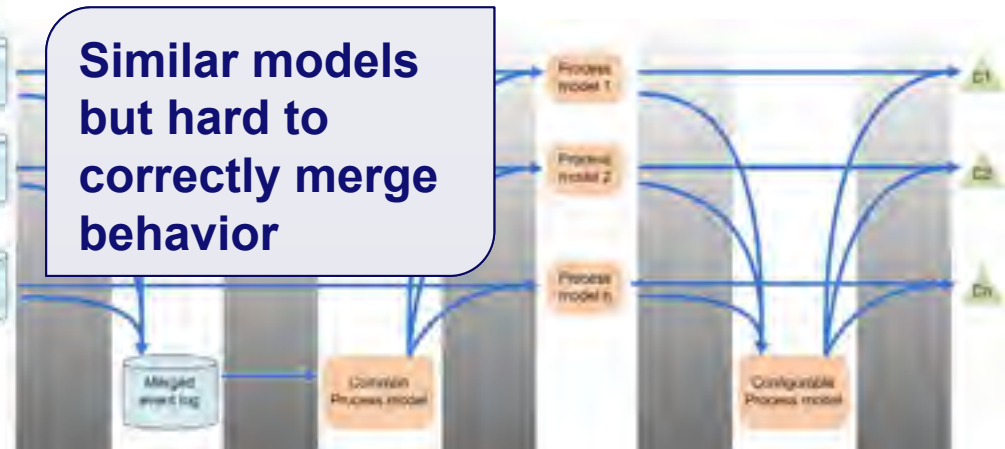
Thank you

Any Questions?

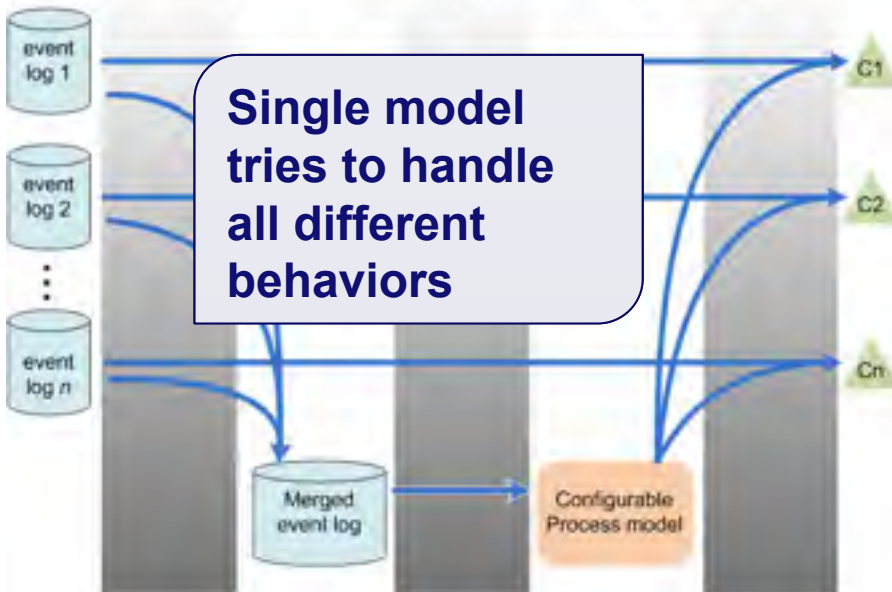
Conclusion



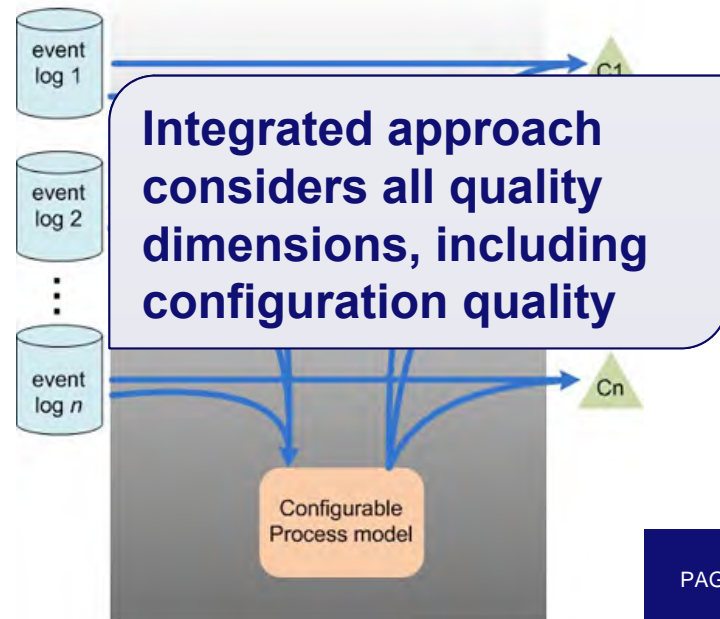
Models are too different to merge



Similar models but hard to correctly merge behavior

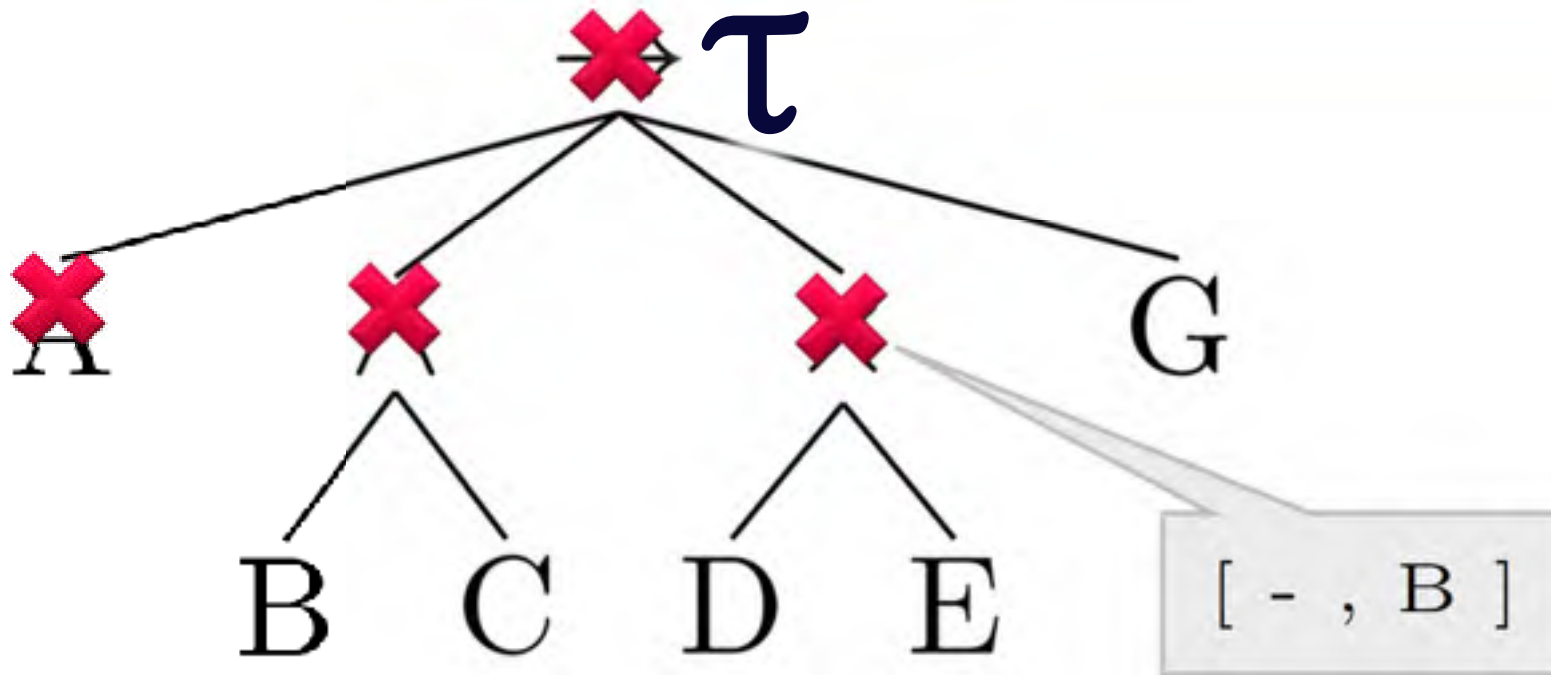


Single model tries to handle all different behaviors

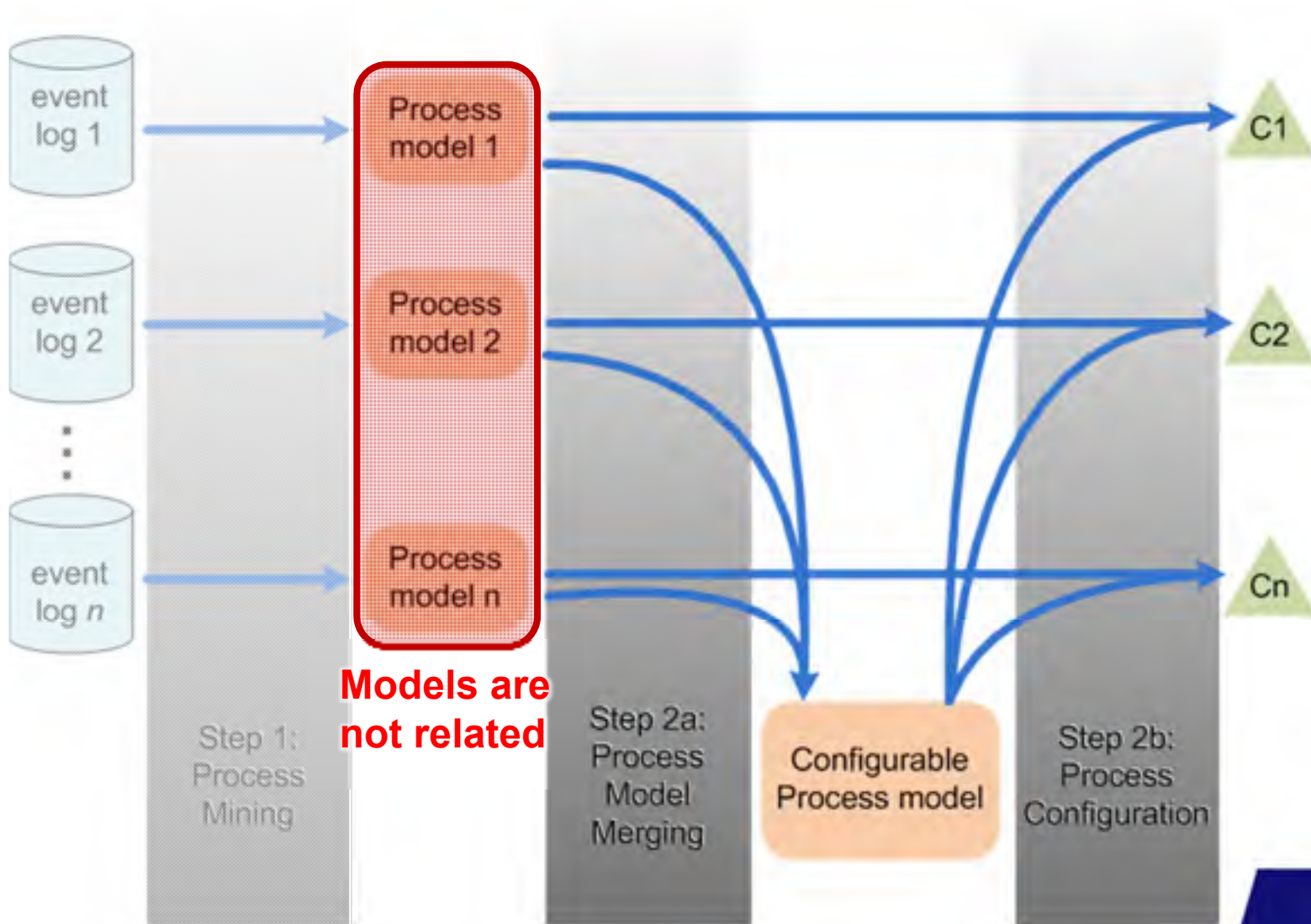


Integrated approach considers all quality dimensions, including configuration quality

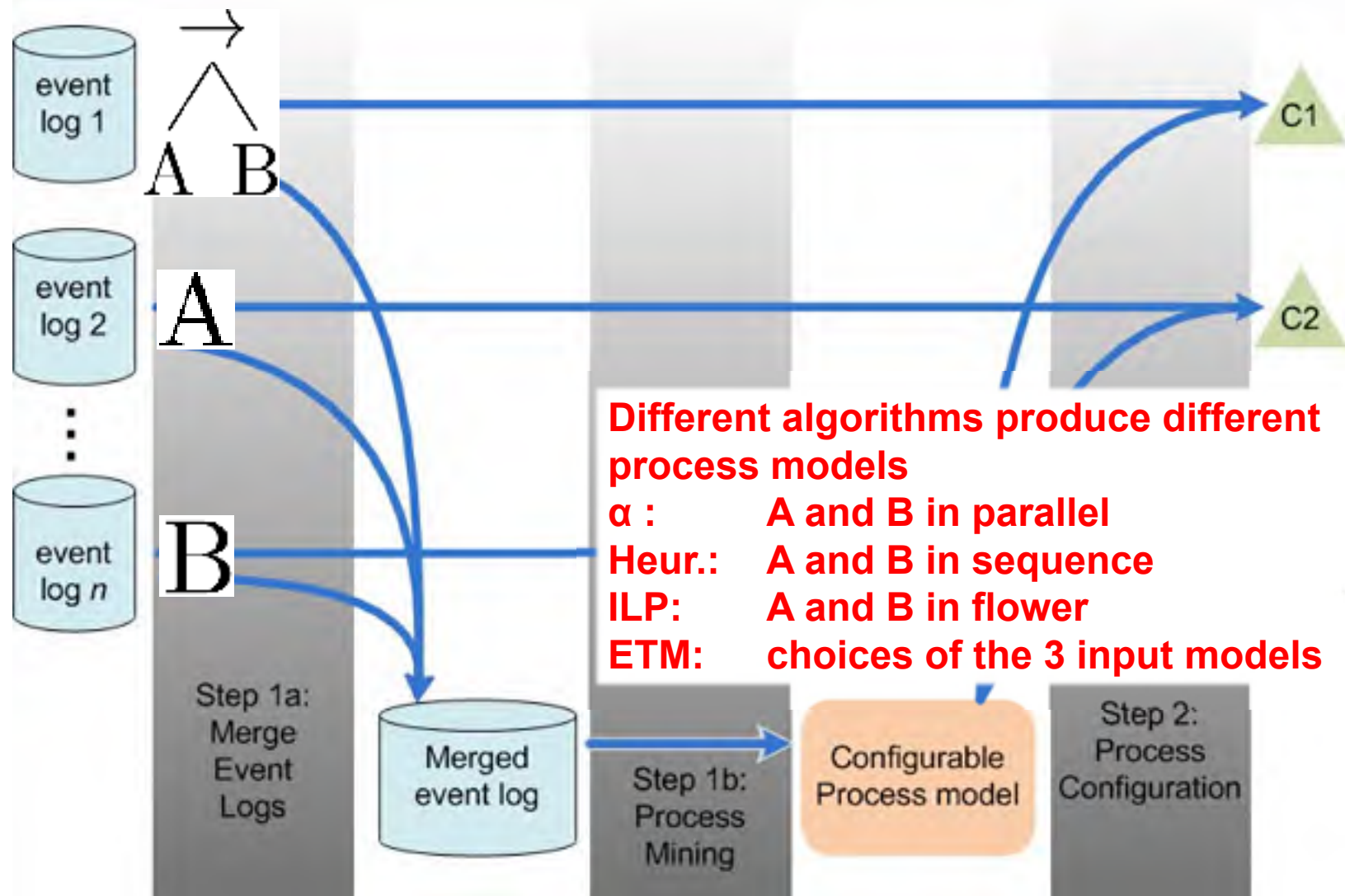
Blocking in Sequence



Issue with Model Merge Approach



Issue with Log Merge Approach



Running Example: 4 Variants

1

#	Trace
38	A B C D F G
26	A B D C F G
12	A B D C E G
8	A B C F G
6	A B C D E G
4	A D C B F G
2	A C D B F G
1	A D B C F G
1	A D B C E G
1	A C B E G
1	A C B F G

B split in 2
Other D variant
Sequential

No D and G

B,C,D in parallel
D can be skipped

B split in 2
Both D variants
No G

2

#	Trace
20	A B1 B2 C D2 E G
50	A B1 B2 C D2 F G

3

#	Trace
120	A C B E
80	A C B F

4

#	Trace
20	A B1 D B2 C E
50	A B1 D2 B2 C F