

Process Mining

COMP9313: Big Data Management

Tutorial / Demo: ProM Framework

The screenshot displays the ProM [5.2] software interface. The main window is titled "outpatientClinicExample.mxml" and features a dashboard with several panels. On the left, a sidebar contains icons for "Dashboard", a funnel, a magnifying glass, and a document. The "Key data" panel lists: Processes (1), Cases (1000), Events (14000), Event classes (18), Event types (2), and Originators (17). The "Events per case" panel shows a green box with statistics: Min 14, Mean 14, and Max 14. The "Event classes per case" panel also shows a green box with the same statistics: Min 14, Mean 14, and Max 14. The "Log info" panel on the right provides details: Source (CPN Tools simulation), Source program (CPN Tools simulation), Start date (2007-01-01 11:00:00), End date (2007-02-21 00:38:00), and Description (CPN Tools simulation log). A button labeled "start analyzing this log" is located at the bottom of the log info panel. The status bar at the bottom indicates "11:17:19 [D] Processing data for buffered log reader completed."

ProM [5.2]

outpatientClinicExample.mxml

outpatientClinicExample.mxml

Dashboard

Key data

Processes	1
Cases	1000
Events	14000
Event classes	18
Event types	2
Originators	17

Events per case

Min 14 Mean 14 Max 14

Event classes per case

Min 14 Mean 14 Max 14

Log info

Source
CPN Tools simulation

Source program
CPN Tools simulation

Start date
2007-01-01 11:00:00

End date
2007-02-21 00:38:00

Description
CPN Tools simulation log

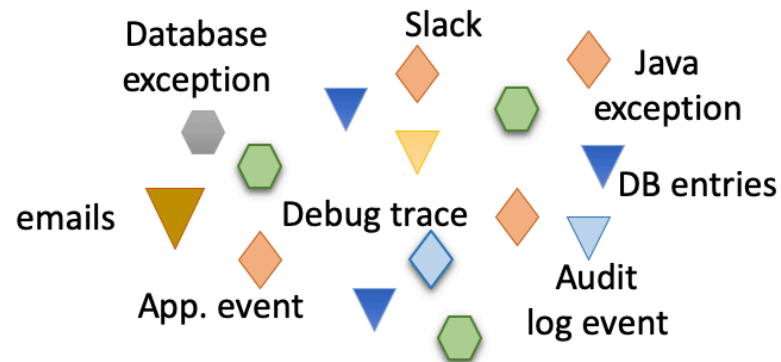
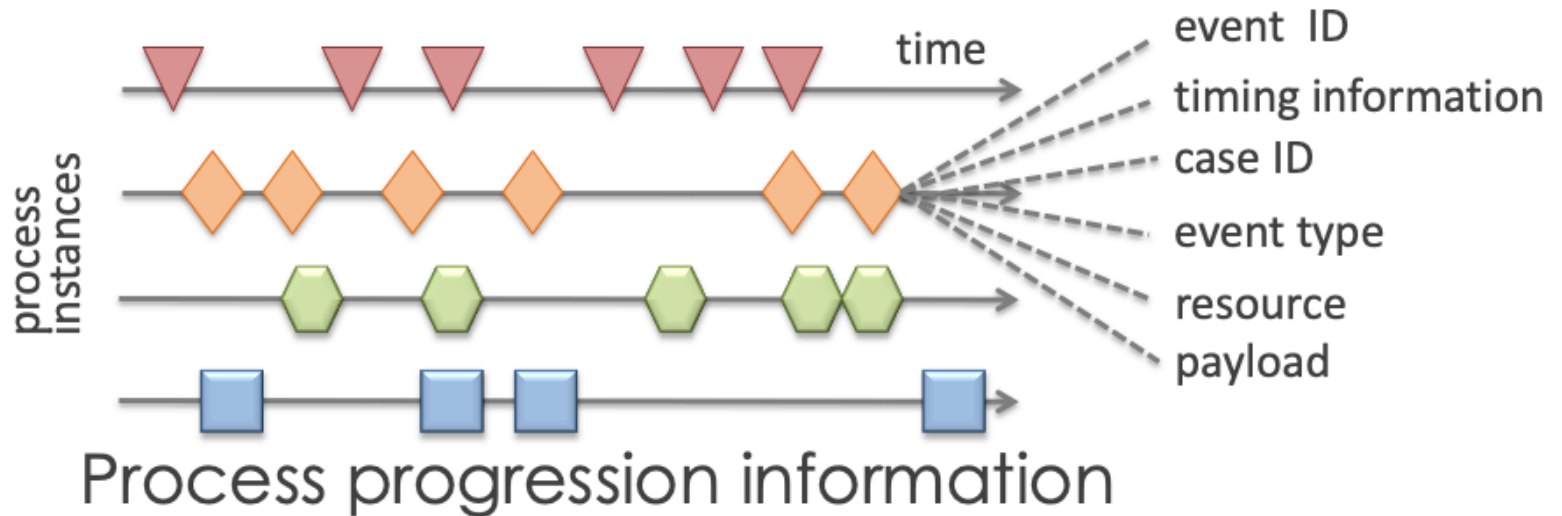
start analyzing this log

11:17:19 [D] Processing data for buffered log reader completed.

Challenges

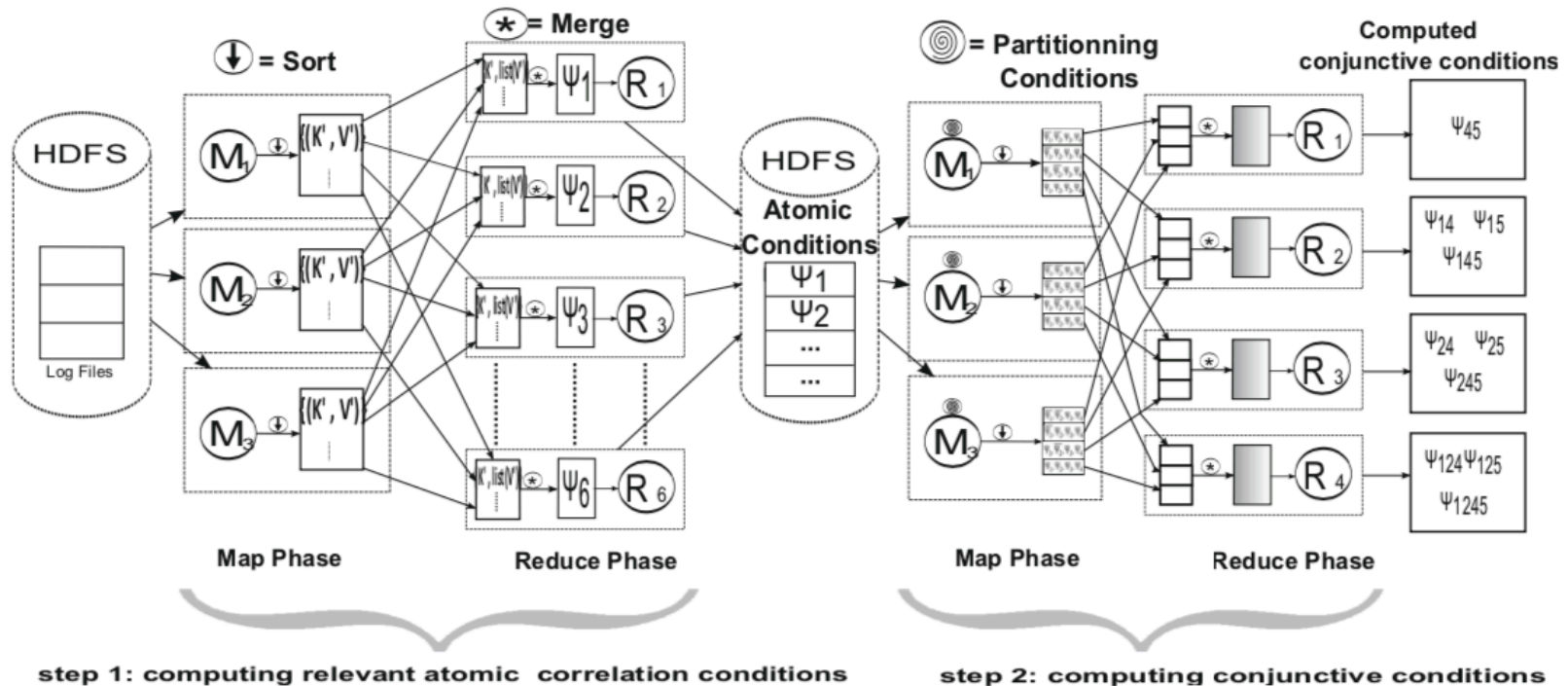
- Huge event logs generated by process-aware information systems (**Volume**)
- Events generated at very high speed (**Velocity**)
- Heterogenous events (**Variety**)
不同的, 异质
- Are the events in the log legit / precise? (**Veracity**)
合法 准确
- Do we get access to all events generated by process-aware information systems? (**Visibility**)
- Are all events really relevant in the context of process mining? (**Value**)

Events and event correlation



What we actually get

Event logs and event correlations



(*) Reguieg et al., 2012, September. Using mapreduce to scale events correlation discovery for business processes mining. In *International Conference on Business Process Management*(pp. 279-284). Springer, Berlin, Heidelberg

Thanks