Event Data in a Relational Database

Classical Log Extraction

To analyze the events using process mining techniques, pick a case identifier for the process, for example the Sales Order id. Any timestamp related to a value of the case identifier becomes an event of that case, sort events by their timestamp.

In case of many-to-many relations between events coming from different tables, two phenomena arise.

Divergence. Events of different instances of the same object are in the same case.

Convergence. Duplication of the same event into different cases.

Convergence leads to false edges and skews statistics. Only invoice B2 was created before its delivery D4, but the model suggest this happened twice.

Divergence leads to more events and flows in the model than actually in the data. There have only been two invoices created, but the model shows three.

Process Mining

A standard process mining technique then returns an end-to-end process across all the business objects touched by the process. In this case the process related to all Sales Orders.

Extracting Artifact Types

We identify clusters of connected tables linked by 1-to-1 relations only. Intuitively, each cluster describes a scheme of similar business objects, within a business object, convergence and divergence cannot arise.

Extracting Artifact Interactions

The 1-to-many and many-to-many relations between clusters and artifact types describe their interactions.

An artifact type may contain event types that are not related. The user may refine the clusters to show only event types that are related.

The user may define the clusters to show only event types that are related.

The user may filter the clusters to show only event types that are related.

The resulting artifact-centric model highlights the main flows across the different objects; object life-cycles can be expanded (here for the Order artifact) or collapsed to focus on specific aspects.

From the artifact-centric model various unusual flows can be identified automatically or by an analyst. Payment topics has been executed too early in some cases. By contrast, the classical process model below has 49% false edges.

Mining Artifact Life-Cycles + Interactions

The artifact type definitions allow for fully automatically extract one event log for each artifact type. From this a life-cycle model for each artifact can be discovered with a classical process mining technique.

Life-cycle model. Describes how each instance of a business object evolves in the process; different instances are separated.