From Complexity to Insight

Querying Large Business Process Models to Improve Quality

PQ 2018
3rd International Workshop on Process Querying
10 September 2018, Sydney, Australia

Kurt Madsen
Process Architect
www.MetaTech.us
About the Author – Kurt Madsen

Professional Experience
• Software Engineer specializing in process analysis and BPM implementation.
• Verticals: banking, manufacturing, health care, aviation, radio, and government.
• Board of directors for WMNF.org, non-profit community conscious radio station in Florida, USA committed to social justice, equality, & creativity.

Education
• Master of science, computer science, New York University, engineering.nyu.edu
• Bachelor of arts in economics, Rutgers University, rutgers.edu
• Taught 20 courses as adjunct instructor at engineering.usf.edu and phoenix.edu

Research Interests
• Extending PQL beyond processes to [TOGAF] enterprise architecture models
• PQL portability across modeling tools (other than ProVision)
• PQL applied to process model drift detection and compliance
Queries → Layers

Models Before Querying and Filtering: Complex and Confusing

Large Process Models
• Multi-year cycle times
• Thousands of workers
• Hundreds of task types
• Dozens of roles
• Complex supply chains

Difficult to Query
• Many stakeholders
• Process maps clouded by irrelevant details
• Poor model searching

→ Complexity + Confusion
Queries → Layers

Models After Querying and Filtering: Clear and Insightful

Query Tools
• Search by task owner
• Search by data flows
• Navigation via composite searches

Filtered Views
• Many stakeholders
• Understanding clouded by irrelevant details
• Poor model searching

→
Clarity + Insight
### Initial Approach to Optimizing Process Time-to-Market

**Lead**

- Design Concept
- Build Prototype

**Lag**

- Build Prototype
- QC Review

#### (A) BPM Workflow Inventory

<table>
<thead>
<tr>
<th>Author/SME (From Activity)</th>
<th>Author/SME (To Activity)</th>
<th>From Activity (Title)</th>
<th>To Activity (Title)</th>
<th>Start (From Activity)</th>
<th>Finish (From Activity)</th>
<th>Start (To Activity)</th>
<th>Finish (To Activity)</th>
<th>Opportunities (manual calc)</th>
<th>Milestone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smithers</td>
<td>Constance</td>
<td>Design concept part</td>
<td>Build prototype part</td>
<td>191</td>
<td>180</td>
<td>185</td>
<td>170</td>
<td>*</td>
<td>5</td>
</tr>
<tr>
<td>Horan</td>
<td>Constance</td>
<td>Safety Tolerances</td>
<td>Build prototype part</td>
<td>185</td>
<td>185</td>
<td>185</td>
<td>170</td>
<td>0</td>
<td>*</td>
</tr>
<tr>
<td>Constance</td>
<td>Jones</td>
<td>Build prototype part</td>
<td>QC part review</td>
<td>185</td>
<td>170</td>
<td>168</td>
<td>167</td>
<td>2</td>
<td>*</td>
</tr>
<tr>
<td>Jones</td>
<td>MacDonald</td>
<td>QC part review</td>
<td>Order 1st batch</td>
<td>168</td>
<td>167</td>
<td>166</td>
<td>160</td>
<td>1</td>
<td>*</td>
</tr>
</tbody>
</table>

#### (B) Gantt Task Inventory

<table>
<thead>
<tr>
<th>ID</th>
<th>Task</th>
<th>Start</th>
<th>Finish</th>
<th>Predecessors</th>
<th>Opportunities (manual calc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Design concept part</td>
<td>191</td>
<td>180</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>Safety Tolerances</td>
<td>185</td>
<td>185</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>QC part review</td>
<td>168</td>
<td>167</td>
<td>5</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>Order 1st batch</td>
<td>166</td>
<td>160</td>
<td>3</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>Build prototype part</td>
<td>185</td>
<td>170</td>
<td>1,2</td>
<td>?</td>
</tr>
</tbody>
</table>

- (A) supports searching and prioritizing workflows
- (B) does not. It requires programmatic iteration

---

© 2002 – 2018
MetaTech, Inc.

From Complexity to Insight:
Querying Large Business Process Models to Improve Quality
Queries → Layers

Step 1: Export BPM model into a Searchable Format for Query

- This research involved querying process exports in CIF.xml and *.XPDL formats. *Future work should move towards real-time queries (inquiry, manipulation, and update) in dynamic, modeling environments.*
- An excerpt of a process model follows. It shows one activity instance of many.

```
01  <activity id="157896" name="CAD-849">
02    <descr>Build prototype car parts</descr>
03    <parent refID="435524"/>
04    <workTime></workTime>
05    <performer refID="467908"/>
06    <customProperties>
07        <property name="Author">
08            <value>John Doe</value>
09        </property>
10    </customProperties>
11  </activity>
```
Queries $\rightarrow$ Layers

Step 2: Manipulate Model Using PQL-like Query Results

- Applying the jQuery below to the previous process model file yields a filtered list of activities which are members of the desired process layer (that satisfies the query in question.).

```xml
01 declare variable $author:="John Doe";
02 for $activity in /process/activities/activity
03 let $activity-id := $activity/@id
04 where $activity/customProperties/property
05 [@name="Author"]/value[matches(., $author)]
06 return <member refID="{$activity-id}" />
```

```xml
01 <modelScenario name="Process layer, author filter">
02   <members>
03     <member refID="157896"/>
04     <member refID="...etc..."/>
05     ...etc...
06   </members>
07 </modelScenario>
```
Step 3: Update Model in Rendering Engine to Show Results

- Applying the jQuery below to the previous process model file yields a filtered list of activities which are members of the desired process layer (that satisfies the query in question.).

```javascript
function highlightWorkflowsByStereotype(model) {
    var bpmParts = model.getComponents();
    for (i = 0; i < bpmParts.length; i++) {
        if (bpmParts[i].getType() == "Workflow") {
            var nextStereotype = bpmParts[i].getStereotype();
            if (nextStereotype == "WfOverlap")
                bpmParts[i].Line.setColor(BLUE);
            else if (nextStereotype == "WfGap")
                bpmParts[i].Line.setColor(GREEN);
            else if (nextStereotype == "WfError")
                bpmParts[i].Line.setColor(RED);
        }
    }
}
```
Queries \to Layers

Models After Querying and Filtering: Clear and Insightful

**Query Tools**
- Search by task owner
- Search by data flows
- Navigation via composite searches

**Filtered Views**
- Many stakeholders
- Understanding clouded by irrelevant details
- Poor model searching

\to

Clarity + Insight
Interconnected Models
Iterative queries into interconnected models enable exploration and discovery within the enterprise.

TOGAF Model Layers
- Strategic
- Organization
- Capabilities
- Processes
- Information
- Applications and Services

Different perspectives
- COO: world-class manufacturing
- CFO: tracing financial data flows
- CTO: legacy migration to cloud
- Arch: integrated knowledge base
- Counsel: regulatory compliance
Filtered list of hyperlinks to models:
- Item 01
- Item 02
- Item 03
- ...etc...

With link syntax as:
```
<a href="https://metatech.us/portfolio.01/viewpoint.02/view.03/item.04.html" target="model_panel">Item.01</a>
```
Filtered list of hyperlinks to models:
- Item 01
- Item 02
- Item 03
- …etc…

With link syntax as:
```html
<a href="https://metatech.us/portfolio.01/viewpoint.02/view.03/item.04.html" target="model_panel">Item.01</a>
```
Filtered list of hyperlinks to models:
- Item 01
- Item 02
- Item 03
- …etc…

With link syntax as:
```html
<a href="https://metatech.us/portfolio.01/viewpoint.02/view.03/item.04.html" target="model_panel">Item.01</a>
```
Filtered list of hyperlinks to models:
- Item 01
- Item 02
- Item 03
- …etc…

With link syntax as:
```
<a href="https://metatech.us/portfolio.01/viewpoint.02/view.03/item.04.html" target="model_panel">Item.01</a>
```
Workflow Data Analysis

Mapping Workflow Data to Cloud-Based Micro-Services

Legacy Forms

Migration to Cloud SaaS

Legacy
- MS Office documents
- Scanned forms
- Archived PDFs
- Manual processes

Migration to cloud
- Retired legacy apps
- Integrated flows
- Service-ready tasks
- Measured progress
Model Drift and Compliance Deviation

Lost Time due to Random Work Arrival

As Is: Worker workload is unpredictable because fraud cases arrive at random.

To Be: Worker utilization is optimized because workload is predictable and efficient.

How it Works: Queue new fraud cases as they arrive. Once enough cases have accumulated, release them in a batch to workers’ To Do lists. (Cases may be released early to meet SLAs.)

Legend
- Available for other work
- Lost Time (due to switching between tasks)
- Productive Time (working on the n\textsuperscript{th} ticket)
Model Drift and Compliance Deviation

Lost Time due to External Dependencies

**As Is:** User polls VISA to see if work is ready.

**Customer Waits**

**To Be:** Fraud investigator is free to work on other tasks and is notified immediately when Visa is ready. Polling tasks are eliminated freeing time for other work.

**Legend**

- Blue = Available for other work
- Red = Lost Time (due to switching between tasks)
- $t_p$ = Productive Time (working on the $n^{th}$ ticket)

**How it works:** Integrate systems with 3rd party (so users don’t have to wait for random events). Then, send push notifications to users when 3rd Party is ready.
Process Query Use Cases at MetaTech

**Querying process logs**
- Not-for-profit, community radio stations lend each other equipment encumbered by donor-imposed usage constraints. Regulations restrict certain usage models.
- Medical device manufacturers release faulty products, get sued, then fined for non-compliance with design specification retention requirements by the gov’t.
- Credit card fraud prevention case workers deviate from process models by rearranging tasks to circumvent dependencies on external partners (e.g., Visa).

**Querying process models**
- Automotive manufacturing process planners struggle to optimize time-to-market due to the scale and complexity of global vehicle development processes.
- Enterprise architects migrating legacy information systems to micro-services in clouds use process querying to extract metadata attached to workflows.
- Medical insurance claims processors use DeMorgan’s laws to re-factor and simplify process steps prior to BPM peer review and quality assurance.
Example PQL Query: Identify Opportunities for Cloud Migration

Goal: replace legacy information systems with cloud-based, micro-services.

1. Identify workflows where process participants used email to hand off information across swim lane boundaries, a practice that leads to rework.

2. Combining the resulting set of workflows with the list of end-of-life systems provided a short-list of migration-eligible systems.

3. Create micro-service interfaces and integrate into BPM modeling tool such that “ServiceReady” Activities are inventoried and available for next-gen modeling.

```
SELECT id FROM workflows AS w WHERE crossesSwimlaneBoundary(w.id) = true AND w.id IN (SELECT id FROM workflows AS w WHERE w.endLink.refId IN (SELECT id FROM activities AS a WHERE has_artifact(a.id) = true AND regExp(a.id, partsDataPattern) = true))
```
From Complexity to Insight:
Querying Large Business Process Models to Improve Quality

Questions?

Kurt Madsen

Telephone: +1 (813) 298-8180
Tampa, Florida, USA
Email: kmadsen@metatech.us
www.MetaTech.us
www.LinkedIn.com/in/KurtMadsen1