Pathfinder: A Relational Query Optimizer Explores XQuery Terrain

Torsten Grust · Jan Rittinger · Jens Teubner · TU München
Pathfinder consumes XQuery with arbitrary expression nestings . . .

- **_literals**: 42, "foo", (), ...
- **arithmetics**: $e_1 + e_2$, $e_1 - e_2$, ...
- **built-in functions**: fn:sum(e), fn:count(e), fn:doc(uri), ...
- **variable bindings**: let $v := e_1$ return $e_2$
- **iteration**: for $v$ at $p$ in $e_1$ return $e_2$
- **conditionals**: if $p$ then $e_1$ else $e_2$
- **sequence construction**: $e_1$, $e_2$
- **user-defd. functions**: $f(e_1, e_2, \ldots, e_n)$
- **element construction**: element $e_1$ { $e_2$ }
- **XPath steps**: $e/\alpha::\nu$ (full axis feature)
Pathfinder consumes XQuery with arbitrary expression nestings . . .

- **_literals**: 42, "foo", (), . . .
- **arithmetics**: $e_1 + e_2$, $e_1 - e_2$, . . .
- **built-in functions**: `fn:sum(e)`, `fn:count(e)`, `fn:doc(uri)`, . . .
- **variable bindings**: `let $v := e_1 return e_2`
- **iteration**: `for $v at $p in e_1 return e_2`
- **conditionals**: `if $p then e_1 else e_2`
- **sequence construction**: $e_1, e_2$
- **user-defd. functions**: $f(e_1, e_2, \ldots, e_n)$
- **element construction**: `element e_1 { e_2 }`
- **XPath steps**: $e/\alpha::\nu$ (full axis feature)

Pathfinder—A Relational Query Optimizer Explores XQuery Terrain
\[\pi\] column projection, renaming
\[\sigma\] row selection
\[\bowtie\] equi-join
\[\times\] Cartesian product
\[\cup, \setminus\] disjoint union, difference
\[\delta\] duplicate elimination
\[\rho\] row numbering
\[\text{staircase join}^1\]
\[\varepsilon, \tau\] element/text node construction\(^1\)
\[\odot\] arithmetic/comparison/Boolean operator \(*\)

\(\blacktriangleright\) Operates on node (not tree!) level and 1NF relations.

\(^1\)Syntactic sugar; expressible by remaining operators.
and compiles them into plans of a standard relational algebra.

\[
\begin{align*}
\pi & \quad \text{column projection, renaming} \\
\sigma & \quad \text{row selection} \\
\bowtie & \quad \text{equi-join} \\
\times & \quad \text{Cartesian product} \\
\cup, \setminus & \quad \text{disjoint union, difference} \\
\delta & \quad \text{duplicate elimination} \\
\rho & \quad \text{row numbering} \\
\Upsilon & \quad \text{staircase join} \\
\varepsilon, \tau & \quad \text{element/text node construction}^1 \\
\odot & \quad \text{arithmetic/comparison/Boolean operator } \ast
\end{align*}
\]

- Operates on node (not tree!) level and 1NF relations.

\[\text{\textsuperscript{1}}\text{Syntactic sugar; expressible by remaining operators.}\]
To combat the resulting plan sizes, Pathfinder uses

Resulting query runtime (MonetDB/XQuery):

112 sec
To combat the resulting plan sizes, Pathfinder uses

1. constant propagation,

Resulting query runtime (MonetDB/XQuery):

103 sec
112 sec
To combat the resulting plan sizes, Pathfinder uses

1. constant propagation,
2. projection pushdown,

Resulting query runtime (MonetDB/XQuery):

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>103 sec</td>
<td>51 sec</td>
<td>112 sec</td>
</tr>
</tbody>
</table>
Relational Query Optimization for XQuery

To combat the resulting plan sizes, Pathfinder uses

1. constant propagation,
2. projection pushdown,
3. functional dependency and data flow analyses, and

Resulting query runtime (MonetDB/XQuery):

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>112 sec</td>
<td>103 sec</td>
<td>51 sec</td>
<td>33 sec</td>
</tr>
</tbody>
</table>
To combat the resulting plan sizes, Pathfinder uses

1. constant propagation,
2. projection pushdown,
3. functional dependency and data flow analyses, and
4. algebraic join detection.

Resulting query runtime (MonetDB/XQuery):

|   | 0.1 sec | 33 sec | 51 sec | 103 sec | 112 sec |
Get your copy of Pathfinder today:

http://www.pathfinder-xquery.org/

Pathfinder is developed at TU München by

- Torsten Grust, Jan Rittinger, and Jens Teubner.

Thanks to our long-term partners

- Peter Boncz, CWI Amsterdam
- Jan Flokstra, U Twente
- Maurice van Keulen, U Twente
- Martin Kersten, CWI Amsterdam
- Stefan Manegold, CWI Amsterdam
- Sjoerd Mullender, CWI Amsterdam
- Ying Zhang, CWI Amsterdam

Pathfinder is supported by the German Research Foundation DFG.