

Evaluating Nested Queries on XML Data

Carlo Sartiani
Dipartimento di Informatica
Università di Pisa

Introduction

- XQuery is a free nesting language
- Nested queries used for
 - reshaping elements
 - creating groups and partitions
 - simulating groupby constructs

Issues in Nested Query Evaluation

- A few nested queries can be unnested
 - the OQL lesson
- OIDs issues
 - XML nodes have OIDs
- position issues
 - XML is ordered

More Issues in Nested Query Evaluation

- Creating new nodes

```
return <ideas> {...} </ideas>
```

- Reshaping existing sub-trees

```
return <name> data($n) </name>
```

- Refreshing OIDs of existing nodes

```
return {$book}
```

Node Taxonomy

- *Natural* nodes: the database
- *Artificial* nodes: database nodes shared in nested query results
- *Synthetic* nodes: newly created nodes

Node Taxonomy

Example

```
for $a in input()/*author
let $title_list :=
  for $b in input()/*,
    $t in $b/title
  where some $ba in $b/author
    satisfies ($ba is $a)
  return <pubtitle> data($t) </pubtitle>
return <ref> { $a, $title_list} </ref>
```

Nested Query Evaluation Approach

- Sharing artificial nodes: no materialization
- update of node labels through hash tables
 - positional labels (structural labels)
 - EIDs

More on Nested Query Evaluation Approach

- Materialization of synthetic nodes
- kept distinct by queryID and invocationID

Conclusions

- A technique for evaluating nested queries on XML data
- scalable
- relatively fast