Jens Dittrich

Applications for Join Algorithms

joins ;-)

INTERSECT, EXCEPT, ...

GROUP BY [HAVING]

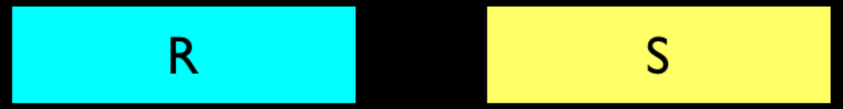
subqueries

Four Principal Classes of Join Algorithms

- nested-loop
- index nested-loop
- hash
- sort-merge

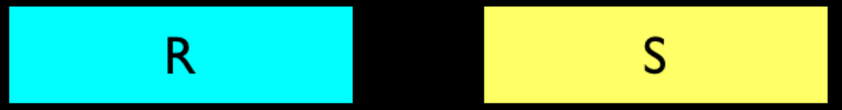
Nested-Loop Join

→ Equi-Join



JP(r,s) := r.x == s.x //definition of the join predicate

Nested-Loop Join



JP(r,s) := r.x == s.x //definition of the join predicate

$JP2(r,s) = r.x \leq s.x$

NestedLoopJoin(R, S, JP(r,s)):

```
ForEach r in R                  //for every tuple in R
    ForEach s in S:             //for every tuple in S
        If JP(r,s):             //check join predicate
            output( (r,s) );    //output join result
```

$\sigma_{JP}(R \times S)$

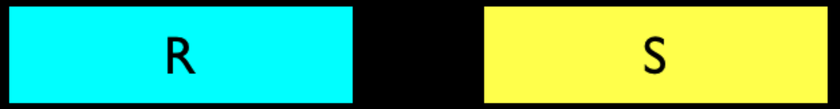
$|R| \times |S|$ comparisons

$|R| = |S| = N$

$\Rightarrow O(N^2)$

Index Nested-Loop Join

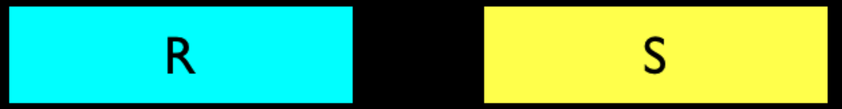
r.a == s.b



JP(r,s) := r.x == s.x //definition of the join predicate

indexOnRX := catalog.get(indexes, R.x); //use existing index on R.x

Index Nested-Loop Join



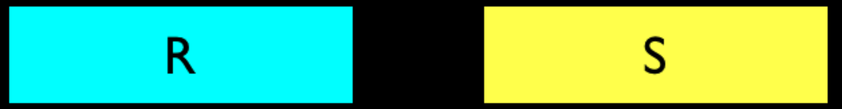

```
JP(r,s) := r.x == s.x                                   //definition of the join predicate
indexOnRX  := catalog.get( indexes, R.x );              //use existing index on R.x
                                                        //precondition for this join algorithm

IndexNestedLoopJoin( indexOnRX, S, JP(r,s) ):
    ForEach s in S:                                     //for every tuple in S
        queryResultSet = indexOnRX.query(s.x);          //query index for this s (aka probe the index)
```

42

↑ point-query

Index Nested-Loop Join




```
JP(r,s) := r.x == s.x                                   //definition of the join predicate
indexOnRX  := catalog.get( indexes, R.x );              //use existing index on R.x
                                                        //precondition for this join algorithm

IndexNestedLoopJoin( indexOnRX, S, JP(r,s) ):
    ForEach s in S:                                     //for every tuple in S
        queryResultSet = indexOnRX.query(s.x);          //query index for this s (aka probe the index)
        If queryResultSet NOT empty:                    //did the query return results?
            output( {s} × queryResultSet );             //output join results
```

e.g. $|queryResultSet| = 7$

$\Rightarrow |\{s\} \times queryResultSet| = 7$