3. Join Ordering

- Basics
- Search Space
- Greedy Heuristics
- IKKBZ
- MVP
- Dynamic Programming
- Simplifying the Query Graph
- Adaptive Optimization
- Generating Permutations
- Transformative Approaches
- Randomized Approaches
- Metaheuristics
- Iterative Dynamic Programming
- Order Preserving Joins
Queries Considered

Concentrate on join ordering, that is:

- conjunctive queries
- simple predicates
- predicates have the form $a_1 = a_2$ where $a_1$ is an attribute and $a_2$ is either an attribute or a constant
- even ignore constants in some algorithms

We join relations $R_1, \ldots, R_n$, where $R_i$ can be

- a base relation
- a base relation including selections
- a more complex building block or access path

Pretending to have a base relation is ok for now.
Query Graph

Queries of this type can be characterized by their query graph:

- the query graph is an undirected graph with $R_1, \ldots, R_n$ as nodes
- a predicate of the form $a_1 = a_2$, where $a_1 \in R_i$ and $a_2 \in R_j$ forms an edge between $R_i$ and $R_j$ labeled with the predicate
- a predicate of the form $a_1 = a_2$, where $a_1 \in R_i$ and $a_2$ is a constant forms a self-edge on $R_i$ labeled with the predicate
- most algorithms will not handle self-edges, they have to be pushed down
Sample Query Graph

student \(\text{sno}=\text{asno}\) attend

\(\bigcap\)

professor \(\text{pno}=\text{lpno}\) lecture

pname="Sokrates"
Shapes of Query Graphs

- chains
- cycles
- stars
- cliques
- cyclic
- tree
- grid

- real world queries are somewhere in-between
- chain, cycle, star and clique are interesting to study
- they represent certain kind of problems and queries
Join Trees

A join tree is a binary tree with

- join operators as inner nodes
- relations as leaf nodes

Algorithms will produce different kinds of join trees

- ordered or unordered
- with cross products or without

The most common case is ordered, without cross products
Shape of Join Trees

Commonly used classes of join trees:

- left-deep tree
- right-deep tree
- zigzag tree
- bushy tree

The first three are summarized as *linear trees*. 
Join Selectivity

Input:
- cardinalities $|R_i|$
- selectivities $f_{i,j}$: if $p_{i,j}$ is the join predicate between $R_i$ and $R_j$, define

$$f_{i,j} = \frac{|R_i \bowtie_{p_{i,j}} R_j|}{|R_i \times R_j|}$$

Calculate:
- result cardinality:

$$|R_i \bowtie_{p_{i,j}} R_j| = f_{i,j}|R_i||R_j|$$

Rational: The selectivity can be computed/estimated easily (ideally).
Cardinality of Join Trees

Given a join tree $T$, the result cardinality $|T|$ can be computed recursively as

$$|T| = \begin{cases} |R_i| & \text{if } T \text{ is a leaf } R_i \\ \left( \prod_{R_i \in T_1, R_j \in T_2} f_{i,j} \right) |T_1||T_2| & \text{if } T = T_1 \Join T_2 \end{cases}$$

- allows for easy calculation of join cardinality
- requires only base cardinalities and selectivities
- assumes independence of the predicates
Sample Statistics

As running example, we use the following statistics:

\[
\begin{align*}
|R_1| &= 10 \\
|R_2| &= 100 \\
|R_3| &= 1000 \\
f_{1,2} &= 0.1 \\
f_{2,3} &= 0.2
\end{align*}
\]

- implies query graph \( R_1 - R_2 - R_3 \)
- assume \( f_{i,j} = 1 \) for all other combinations
A Basic Cost Function

Given a join tree $T$, the cost function $C_{out}$ is defined as

$$C_{out}(T) = \begin{cases} 
0 & \text{if } T \text{ is a leaf } R_i \\
|T| + C_{out}(T_1) + C_{out}(T_2) & \text{if } T = T_1 \bowtie T_2
\end{cases}$$

- sums up the sizes of the (intermediate) results
- rational: larger intermediate results cause more work
- we ignore the costs of single relations as they have to be read anyway
**Basic Join Specific Cost Functions**

For single joins:

\[
C_{nlj}(e_1 \Join e_2) = |e_1||e_2|
\]

\[
C_{hj}(e_1 \Join e_2) = 1.2|e_1|
\]

\[
C_{smj}(e_1 \Join e_2) = |e_1| \log(|e_1|) + |e_2| \log(|e_2|)
\]

For sequences of join operators \( s = s_1 \Join \ldots \Join s_n \):

\[
C_{nlj}(s) = \sum_{i=2}^{n} |s_1 \Join \ldots \Join s_{i-1}| |s_i|
\]

\[
C_{hj}(s) = \sum_{i=2}^{n} 1.2|s_1 \Join \ldots \Join s_{i-1}|
\]

\[
C_{smj}(s) = \sum_{i=2}^{n} |s_1 \Join \ldots \Join s_{i-1}| \log(|s_1 \Join \ldots \Join s_{i-1}|) + \sum_{i=2}^{n} |s_i| \log(|s_i|)
\]
Remarks on the Basic Cost Functions

- cost functions are simplistic
- algorithms are modelled very simplified (e.g. 1.2, no n-way sort etc.)
- designed for left-deep trees
- $C_{hj}$ and $C_{smj}$ do not work for cross products (fix: take output cardinality then, which is $C_{nl}$)
- in reality: other parameters than cardinality play a role
- cost functions assume the same join algorithm for the whole join tree
## Sample Cost Calculations

<table>
<thead>
<tr>
<th></th>
<th>$C_{out}$</th>
<th>$C_{nl}$</th>
<th>$C_{hj}$</th>
<th>$C_{smj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1 \Join R_2$</td>
<td>100</td>
<td>1000</td>
<td>12</td>
<td>697.61</td>
</tr>
<tr>
<td>$R_2 \Join R_3$</td>
<td>20000</td>
<td>100000</td>
<td>120</td>
<td>10630.26</td>
</tr>
<tr>
<td>$R_1 \times R_3$</td>
<td>10000</td>
<td>10000</td>
<td>10000</td>
<td>10000.00</td>
</tr>
<tr>
<td>$(R_1 \Join R_2) \Join R_3$</td>
<td>20100</td>
<td>101000</td>
<td>132</td>
<td>11327.86</td>
</tr>
<tr>
<td>$(R_2 \Join R_3) \Join R_1$</td>
<td>40000</td>
<td>300000</td>
<td>24120</td>
<td>32595.00</td>
</tr>
<tr>
<td>$(R_1 \times R_3) \Join R_2$</td>
<td>30000</td>
<td>1010000</td>
<td>22000</td>
<td>143542.00</td>
</tr>
</tbody>
</table>

- costs differ vastly between join trees
- different cost functions result in different costs
- the cheapest plan is always the same here, but relative order varies
- join trees with cross products are expensive
- join order is essential under all cost functions
More Examples

For the query \(|R_1| = 1000, |R_2| = 2, |R_3| = 2, f_{1,2} = 0.1, f_{1,3} = 0.1\) we have costs:

<table>
<thead>
<tr>
<th>Operation</th>
<th>(C_{out})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R_1 \bowtie R_2)</td>
<td>200</td>
</tr>
<tr>
<td>(R_2 \times R_3)</td>
<td>4</td>
</tr>
<tr>
<td>(R_1 \bowtie R_3)</td>
<td>200</td>
</tr>
<tr>
<td>((R_1 \bowtie R_2) \bowtie R_3)</td>
<td>240</td>
</tr>
<tr>
<td>((R_2 \times R_3) \bowtie R_1)</td>
<td>44</td>
</tr>
<tr>
<td>((R_1 \bowtie R_3) \bowtie R_2)</td>
<td>240</td>
</tr>
</tbody>
</table>

- here cross product is best
- but relies on the small sizes of \(|R_2|\) and \(|R_3|\)
- attractive if the cardinality of one relation is small
More Examples (2)

For the query $|R_1| = 10, |R_2| = 20, |R_3| = 20, |R_4| = 10$, $f_{1,2} = 0.01$, $f_{2,3} = 0.5$, $f_{3,4} = 0.01$

we have costs:

<table>
<thead>
<tr>
<th></th>
<th>$C_{out}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1 \bowtie R_2$</td>
<td>2</td>
</tr>
<tr>
<td>$R_2 \bowtie R_3$</td>
<td>200</td>
</tr>
<tr>
<td>$R_3 \bowtie R_4$</td>
<td>2</td>
</tr>
<tr>
<td>$((R_1 \bowtie R_2) \bowtie R_3) \bowtie R_4$</td>
<td>24</td>
</tr>
<tr>
<td>$((R_2 \times R_3) \bowtie R_1) \bowtie R_4$</td>
<td>222</td>
</tr>
<tr>
<td>$(R_1 \bowtie R_2) \bowtie (R_3 \bowtie R_4)$</td>
<td>6</td>
</tr>
</tbody>
</table>

- covers all join trees due to the symmetry of the query
- the bushy tree is better than all join trees
Symmetry and ASI

- cost function $C_{impl}$ is called symmetric if 
  \[ C_{impl}(e_1 \bowtie^{impl} e_2) = C_{impl}(e_2 \bowtie^{impl} e_1) \]
- for symmetric cost functions commutativity can be ignored
- ASI: *adjacent sequence interchange* (see IKKBZ algorithm for a definition)

Our basic cost functions can be classified as:

<table>
<thead>
<tr>
<th></th>
<th>ASI</th>
<th>$\neg$ASI</th>
</tr>
</thead>
<tbody>
<tr>
<td>symmetric</td>
<td>$C_{out}$</td>
<td>$C_{smj}$</td>
</tr>
<tr>
<td>$\neg$symmetric</td>
<td>$C_{hj}$</td>
<td>-</td>
</tr>
</tbody>
</table>

- more complex cost functions are usually $\neg$ASI, often also $\neg$symmetric
- symmetry and especially ASI can be exploited during optimization
Classification of Join Ordering Problems

We distinguish four different dimensions:

1. query graph class: *chain*, *cycle*, *star*, and *clique*
2. join tree structure: *left-deep*, *zig-zag*, or *bushy* trees
3. join construction: *with* or *without* cross products
4. cost function: *with* or *without* ASI property

In total, 48 different join ordering problems.
Reminder: Catalan Numbers

The number of binary trees with $n$ leave nodes is given by $C(n - 1)$, where $C(n)$ is defined as

$$C(n) = \begin{cases} 1 & \text{if } n = 0 \\ \sum_{k=0}^{n-1} C(k)C(n - k - 1) & \text{if } n > 0 \end{cases}$$

It can be written in a closed form as

$$C(n) = \frac{1}{n + 1} \binom{2n}{n}$$

The Catalan Numbers grown in the order of $\Theta(4^n/n^{3/2})$
### Number Of Join Trees with Cross Products

<table>
<thead>
<tr>
<th>Type</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>left deep</td>
<td>$n!$</td>
</tr>
<tr>
<td>right deep</td>
<td>$n!$</td>
</tr>
<tr>
<td>zig-zag</td>
<td>$n!2^{n-2}$</td>
</tr>
<tr>
<td>bushy</td>
<td>$n!C(n-1)$</td>
</tr>
</tbody>
</table>

$$= \frac{(2n-2)!}{(n-1)!}$$

- rational: number of leaf combinations ($n!$) $\times$ number of unlabeled trees (varies)
- grows exponentially
- increases even more with a flexible tree structure
Chain Queries, no Cross Products

Let us denote the number of left-deep join trees for a chain query $R_1 - \ldots - R_n$ as $f(n)$

- obviously $f(0) = 1, f(1) = 1$
- for $n > 1$, consider adding $R_n$ to all join trees for $R_1 - \ldots - R_{n-1}$
- $R_n$ can be added at any position following $R_{n-1}$
- let's denote the position of $R_{n-1}$ from the bottom with $k$ ($[1, n-1]$)
- there are $n - k$ join trees for adding $R_n$ after $R_{n-1}$
- one additional tree if $k = 1$, $R_n$ can also be added before $R_{n-1}$
- for $R_{n-1}$ to be at $k$, $R_{n-k} - \ldots R_{n-2}$ must be below it. $f(k - 1)$ trees

for $n > 1$:

$$f(n) = 1 + \sum_{k=1}^{n-1} f(k - 1) \times (n - k)$$
Chain Queries, no Cross Products (2)

The number of left-deep join trees for chain queries of size $n$ is

$$ f(n) = \begin{cases} 
1 & \text{if } n < 2 \\
1 + \sum_{k=1}^{n-1} f(k-1) \times (n-k) & \text{if } n \geq 2
\end{cases} $$

solving the recurrence gives the closed form

$$ f(n) = 2^{n-1} $$

- generalization to zig-zag as before
Chain Queries, no Cross Products (3)

The generalization to bushy trees is not as obvious

• each subtree must contain a subchain to avoid cross products
• thus do not add single relations but subchains
• whole chain must be $R_1 - \ldots - R_n$, cut anywhere
• consider commutativity (two possibilities)

This leads to the formula

$$f(n) = \begin{cases} 
1 & \text{if } n < 2 \\
\sum_{k=1}^{n-1} 2f(k)f(n-k) & \text{if } n \geq 2
\end{cases}$$

solving the recurrence gives the closed form

$$f(n) = 2^{n-1}C(n-1)$$
Star Queries, no Cross Products

Consider a star query with $R_1$ at the center and $R_2, \ldots, R_n$ as satellites.

- the first join must involve $R_1$
- afterwards all other relations can be added arbitrarily

This leads to the following formulas:

- left-deep: $2 \times (n - 1)!$
- zig-zag: $2 \times (n - 1)! \times 2^{n-2} = (n - 1)! \times 2^{n-1}$
- bushy: no bushy trees possible ($R_1$ required), same as zig-zag
Clique Queries, no Cross Products

- in a clique query, every relation is connected to each other
- thus no join tree contains cross products
- all join trees are valid join trees, the number is the same as with cross products
## Sample Numbers, without Cross Products

<table>
<thead>
<tr>
<th>n</th>
<th>Chain Queries</th>
<th>Star Queries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left-Deep</td>
<td>Zig-Zag</td>
</tr>
<tr>
<td></td>
<td>$2^{n-1}$</td>
<td>$2^{2n-3}$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>128</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>512</td>
</tr>
<tr>
<td>7</td>
<td>64</td>
<td>2048</td>
</tr>
<tr>
<td>8</td>
<td>128</td>
<td>8192</td>
</tr>
<tr>
<td>9</td>
<td>256</td>
<td>32768</td>
</tr>
<tr>
<td>10</td>
<td>512</td>
<td>131072</td>
</tr>
</tbody>
</table>
Sample Numbers, with Cross Products

<table>
<thead>
<tr>
<th>n</th>
<th>Left-Deep</th>
<th>Zig-Zag $n!2^{n-2}$</th>
<th>Bushy $n!C(n-1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$n!$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>96</td>
<td>120</td>
</tr>
<tr>
<td>5</td>
<td>120</td>
<td>960</td>
<td>1680</td>
</tr>
<tr>
<td>6</td>
<td>720</td>
<td>11520</td>
<td>30240</td>
</tr>
<tr>
<td>7</td>
<td>5040</td>
<td>161280</td>
<td>665280</td>
</tr>
<tr>
<td>8</td>
<td>40320</td>
<td>2580480</td>
<td>17297280</td>
</tr>
<tr>
<td>9</td>
<td>362880</td>
<td>46448640</td>
<td>518918400</td>
</tr>
<tr>
<td>10</td>
<td>3628800</td>
<td>968972800</td>
<td>17643225600</td>
</tr>
</tbody>
</table>
## Problem Complexity

<table>
<thead>
<tr>
<th>query graph</th>
<th>join tree</th>
<th>cross products</th>
<th>cost function</th>
<th>complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>general</td>
<td>left-deep</td>
<td>no</td>
<td>ASI</td>
<td>NP-hard</td>
</tr>
<tr>
<td>tree/star/chain</td>
<td>left-deep</td>
<td>no</td>
<td>ASI, 1 joint. NLJ+SMJ</td>
<td>P, NP-hard</td>
</tr>
<tr>
<td>star</td>
<td>left-deep</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>general/tree/star</td>
<td>left-deep</td>
<td>yes</td>
<td>ASI</td>
<td>NP-hard open</td>
</tr>
<tr>
<td>chain</td>
<td>left-deep</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>general tree</td>
<td>bushy</td>
<td>no</td>
<td>ASI</td>
<td>NP-hard open</td>
</tr>
<tr>
<td>star</td>
<td>bushy</td>
<td>no</td>
<td>-</td>
<td>P</td>
</tr>
<tr>
<td>chain</td>
<td>bushy</td>
<td>no</td>
<td>any</td>
<td>P</td>
</tr>
<tr>
<td>general tree/star</td>
<td>bushy</td>
<td>no</td>
<td>ASI</td>
<td>NP-hard</td>
</tr>
<tr>
<td>chain</td>
<td>bushy</td>
<td>yes</td>
<td>ASI</td>
<td>NP-hard</td>
</tr>
<tr>
<td>general tree/star</td>
<td>bushy</td>
<td>yes</td>
<td>ASI</td>
<td>NP-hard</td>
</tr>
</tbody>
</table>
Greedy Heuristics - First Algorithm

- search space of joins trees is very large
- greedy heuristics produce suitable join trees very fast
- suitable for large queries

For the first algorithm we consider:

- left-deep trees
- no cross products
- relations ordered to some weight function (e.g. cardinality)

Note: the algorithms produces a sequence of relations; it uniquely identifies the left-deep join tree.
Greedy Heuristics - First Algorithm (2)

GreedyJoinOrdering-1\( (R = \{ R_1, \ldots, R_n \}, w : R \rightarrow \mathbb{R}) \)

**Input:** a set of relations to be joined and weight function

**Output:** a join order

\[ S = \epsilon \]

while \(|R| > 0\) {

\[ m = \arg \min_{R_i \in R} w(R_i) \]

\[ R = R \setminus \{ m \} \]

\[ S = S \circ \langle m \rangle \]

}

return \( S \)

• disadvantage: fixed weight functions

• already chosen relations do not affect the weight

• e.g. does not support minimizing the intermediate result
Greedy Heuristics - Second Algorithm

GreedyJoinOrdering-2\((R = \{R_1, \ldots, R_n\}, w : R, R^* \rightarrow \mathbb{R})\)

**Input:** a set of relations to be joined and weight function

**Output:** a join order

\(S = \epsilon\)

while \(|R| > 0\) {

\(m = \arg\min_{R_i \in R} w(R_i, S)\)

\(R = R \setminus \{m\}\)

\(S = S \circ <m>\)

}

return \(S\)

- can compute relative weights
- but first relation has a huge effect
- and the fewest information available
Greedy Heuristics - Third Algorithm

GreedyJoinOrdering-3\( (R = \{R_1, \ldots, R_n\}, w : R, R^* \rightarrow \mathbb{R}) \)

**Input:** a set of relations to be joined and weight function

**Output:** a join order

\( S = \emptyset \)

for each \( R_i \in R \) {

\( R' = R \setminus \{R_i\} \)

\( S' = < R_i > \)

while (\(|R'| > 0\)) {

\( m = \arg \min_{R_j \in R'} w(R_j, S') \)

\( R' = R' \setminus \{m\} \)

\( S' = S' \circ < m > \)

}

\( S = S \cup \{S'\} \)

return \( \arg \min_{S' \in S} w(S'[n], S'[1 : n - 1]) \)

- commonly used: minimize selectivities (MinSel)
Greedy Operator Ordering

- the previous greedy algorithms only construct left-deep trees
- Greedy Operator Ordering (GOO) [1] constructs bushy trees

Idea:

- all relations have to be joined somewhere
- but joins can also happen between whole join trees
- we therefore greedily combine join trees (which can be relations)
- combine join trees such that the intermediate result is minimal
Greedy Operator Ordering (2)

GOO($R = \{R_1, \ldots, R_n\}$)

**Input:** a set of relations to be joined

**Output:** a join tree

$T = R$

while $|T| > 1$

$(T_i, T_j) = \arg\min_{(T_i \in T, T_j \in T), T_i \neq T_j} |T_i \Join T_j|$

$T = (T \setminus \{T_i\}) \setminus \{T_j\}$

$T = T \cup \{T_i \Join T_j\}$

} 

return $T_0 \in T$

- constructs the result bottom up
- join trees are combined into larger join trees
- chooses the pair with the minimal intermediate result in each pass
IKKBZ

Polynomial algorithm for join ordering (original [2], improved [3])

- produces optimal left-deep trees without cross products
- requires acyclic join graphs
- cost function must have ASI property
- join method must be fixed

Can be used as heuristic if the requirements are violated
Overview

- the algorithms considers each relation as first relation to be joined
- it tries to order the other relations by "benefit" (rank)
- if the ordering violates the query constraints, it constructs compounds
- the compounds guarantee the constraints (locally) and are again ordered by benefit
- related to a known job-ordering algorithm
Cost Function

The IKKBZ algorithm considers only cost functions of the form

\[ C(T_i \bowtie R_j) = |T_i| \cdot h_j(|R_j|) \]

- each relation \( R_j \) can have its own \( h_j \)
- we denote the set of \( h_j \) by \( H \), writing \( C_H \) for the parametrized cost function
- examples: \( h_j \equiv 1.2 \) for \( C_{hj} \), \( h_j \equiv id \) for \( C_{nl} \)

We will often use cardinalities, thus we define \( n_i \):

- \( n_i \) is the cardinality of \( R_i \) (\( n = |R_i| \))
- \( h_i(n_i) \) is are the costs per input tuple of a join with \( R_i \)
Precedence Graph

Given a query graph $G = (V, E)$ and a starting relation $R_k$, we construct the directed precedence graph $G^P_k = (V^P_k, E^P_k)$ rooted in $R_k$ as follows:

1. choose $R_k$ as the root node of $G^P_k$, $V^P_k = \{ R_k \}$

2. while $|V^P_k| < |V|$, choose a $R_i \in V \setminus V^P_k$ such that
   $\exists R_j \in V^P_k : (R_j, R_i) \in E$. Add $R_i$ to $V^P_k$ and $R_j \rightarrow R_i$ to $E^P_k$.

The precedence graph describes the (partial) ordering of joins implied by the query graph.
Sample Precedence Graph

query graph

precedence graph rooted in $R_1$
Conformance to a Precedence Graph

A sequence $S = v_1, \ldots, v_k$ of nodes conforms to a precedence graph $G = (V, E)$ if the following conditions are satisfied:

1. $\forall i \in [2, k] \exists j \in [1, i]: (v_j, v_i) \in E$
2. $\forall i \in [1, k], j \in [i, k]: (v_j, v_i) \in E$

Note: IKKBZ constructs left-deep trees, therefore it is sufficient to consider sequences.
Notations

For non-empty sequences $S_1$ and $S_2$ and a precedence graph $G = (V, E)$, we write $S_1 \rightarrow S_2$ if $S_1$ must occur before $S_2$. More precisely $S_1 \rightarrow S_2$ iff:

1. $S_1$ and $S_2$ conform to $G$
2. $S_1 \cap S_2 = \emptyset$
3. $\exists v_i, v_j \in V : v_i \in S_1 \land v_j \in S_2 \land (v_i, v_j) \in E$
4. $\forall v_i, v_j \in V : v_i \in S_1 \land v_j \in V \setminus S_1 \setminus S_2 \land (v_i, v_j) \in E$

Further, we write

$$R_{1,2,\ldots,k} = R_1 \Join R_2 \Join \ldots \Join R_k$$

$$n_{1,2,\ldots,k} = |R_{1,2,\ldots,k}|$$
Selectivities

For a given precedence graph, let $R_i$ be a relation and $\mathcal{R}_i$ be the set of relations from which there exists a path to $R_i$

- in any conforming join tree which includes $R_i$, all relations from $\mathcal{R}_i$ must be joined first
- all other relations $R_j$ that might be joined before $R_i$ will have no connection to $R_i$, thus $f_{i,j} = 1$

Hence, we can define the selectivity of the join with $R_i$ as

$$s_i = \begin{cases} 
1 & \text{if } |\mathcal{R}_i| = 0 \\
\prod_{R_j \in \mathcal{R}_i} f_{i,j} & \text{if } |\mathcal{R}_i| > 0 
\end{cases}$$

Note: we call the $s_i$ a selectivities, although they depend on the precedence graph
Cardinalities

If the query graph is a chain (totally ordered), the following conditions holds:

\[ n_{1,2,\ldots,k} = s_k \cdot |R_k| \cdot |R_{1,2,\ldots,k-1}| = |s_k| \cdot n_k \cdot n_{1,2,\ldots,k-1} \]

As a closed form, we can write

\[ n_{1,2,\ldots,k} = \prod_{i=1}^{k} s_i n_i \]

as \( s_1 = 1 \)
Costs

The costs for a totally ordered precedence graph $G$ can be computed as follows:

$$C_H(G) = \sum_{i=2}^{n} [n_{1,2,...,i-1} h_i(n_i)]$$

$$= \sum_{i=2}^{n} [(\prod_{j=1}^{i} s_jn_j) h_i(n_i)]$$

- if we choose $h_i(n_i) = s_i n_i$ then $C_H \equiv C_{out}$
- the factor $s_i n_i$ determines how much the input relation to be joined with $R_i$ changes its cardinality after the join has been performed
- if $s_i n_i$ is less than one, we call the join *decreasing*, if it is larger than one, we call the join *increasing*
Costs (2)

For the algorithm, we prefer a (equivalent) recursive definition of the cost function:

\[
\begin{align*}
C_H(\epsilon) &= 0 \\
C_H(R_i) &= 0 \text{ if } R_i \text{ is the root} \\
C_H(R_i) &= h_i(n_i) \text{ else} \\
C_H(S_1S_2) &= C_H(S_1) + T(S_1) \ast C_H(S_2)
\end{align*}
\]

where

\[
\begin{align*}
T(\epsilon) &= 1 \\
T(S) &= \prod_{R_i \in S} s_i n_i
\end{align*}
\]
**ASI Property**

Let $A$ and $B$ be two sequences and $V$ and $U$ two non-empty sequences. We say a cost function $C$ has the *adjacent sequence interchange property* (ASI property), if and only if there exists a function $T$ and a rank function defined as

$$
rank(S) = \frac{T(S) - 1}{C(S)}
$$

such that the following holds

$$
C(AUVB) \leq C(AVUB) \iff rank(U) \leq rank(V)
$$

if $AUVB$ and $AVUB$ satisfy the precedence constraints imposed by a given precedence graph.
First Lemma

**Lemma:** The cost function $C_h$ has the ASI-Property.

**Proof:** The proof can be derived from the definition of $C_H$:

$$C_H(AUVB) = C_H(A) + T(A)C_H(U) + T(A)T(U)C_H(V) + T(A)T(U)T(V)C_H(B)$$

and, hence,

$$C_H(AUVB) - C_H(AVUB) = T(A)[C_H(V)(T(U) - 1) - C_H(U)(T(V) - 1)]$$

$$= T(A)C_H(U)C_H(V)[\text{rank}(U) - \text{rank}(V)]$$

The lemma follows.
Module

Let $M = \{A_1, \ldots, A_n\}$ be a set of sequences of nodes in a given precedence graph. Then, $M$ is called a module, if for all sequences $B$ that do not overlap with the sequences in $M$, one of the following conditions holds:

- $B \rightarrow A_i$, $\forall A_i \in M$
- $A_i \rightarrow B$, $\forall A_i \in M$
- $B \not\rightarrow A_i$ and $A_i \not\rightarrow B$, $\forall A_i \in M$
Second Lemma

**Lemma:** Let $C$ be any cost function with the ASI property and $\{A, B\}$ a module. If $A \rightarrow B$ and additional $\text{rank}(B) \leq \text{rank}(A)$, then we find an optimal sequence among those in which $B$ directly follows $A$.

**Proof:** by contradiction. Every optimal permutation must have the form $UAVBW$ since $A \rightarrow B$.

Assumption: $V \neq \epsilon$ for all optimal solutions.

- if $\text{rank}(V) \leq \text{rank}(A)$, we can exchange $V$ and $A$ without increasing the costs.
- if $\text{rank}(A) \leq \text{rank}(V)$, $\text{rank}(B) \leq \text{rank}(V)$ due to the transitivity of $\leq$. Hence, we can exchange $B$ and $V$ without increasing the costs.

Both exchanges produces legal sequences since $\{A, B\}$ is a module.
Contradictory Sequences and Compound Relations

- if the precedence graph demands $A \rightarrow B$ but $rank(B) \leq rank(A)$, we speak of *contradictory sequences* $A$ and $B$
- second lemma $\Rightarrow$ no non-empty subsequence can occur between $A$ and $B$
- we combine $A$ and $B$ into a new single node replacing $A$ and $B$
- this nodes represents a *compound relation* comprising of all relations in $A$ and $B$
- its cardinality is computed by multiplying the cardinalities of all relations in $A$ and $B$
- its selectivity is the product of all selectivities $s_i$ of relations $R_i$ contained in $A$ and $B$
Normalization and Denormalization

- the continued process of building compound relations until no more contradictory sequences exist is called *normalization*
- the opposite step, replacing a compound relation by the sequence of relations it was derived from is called *denormalization*
Algorithm

IKKBZ(G, CH)

**Input:** an acyclic query graph G for relations \( R = \{R_1, \ldots, R_n\} \),
a cost function \( CH \)

**Output:** the optimal left-deep tree

\( S = \emptyset \)

**for each** \( R_i \in R \) {

\( G_i = \) the precedence graph derived from \( G \) rooted at \( R_i \)

\( S_i = \text{IKKBZ-Sub}(G_i, CH) \)

\( S = S \cup \{S_i\} \)

}

**return** \( \arg \min_{S_i \in S} CH(S_i) \)

- considers each relation as starting relation
- constructs the precedence graph and starts the main algorithm
Algorithm (2)

IKKBZ-Sub($G_i, C_H$)

**Input:** a precedence graph $G_i$ for relations $R = \{R_1, \ldots, R_n\}$ rooted at $R_i$, a cost function $C_H$

**Output:** the optimal left-deep tree under $G_i$

**while** $G_i$ is not a chain {
    $r =$ a subtree of $G_i$ whose subtrees are chains
    IKKBZ-Normalize($r$)
    merge the chains under $r$ according to the rank function (ascending)
}

IKKBZ-Denormalize($G_i$)

**return** $G_i$

- transforms the precedence graph into a chain
- wherever there are multiple choices, there are serialized according to the rank
- normalization required to preserve the query graph
Algorithm (3)

IKKBZ-Normalize($R$)

Input: a subtree $R$ of a precedence graph $G = (V, E)$

Output: a normalized subtree

while $\exists r, c \in T, (r, c) \in E : rank(r) > rank(c)$ {
    replace $r$ and $c$ by a compound relation $r'$ that represent $rc$
}

return $R$

- merges relations that would have been reorder if only considering the rank
- guarantees that the rank is ascending in each subchain
Algorithm (4)

IKKBZ-Denormalize($R$)

**Input:** a precedence graph $R$ containing relations and compound relations

**Output:** a denormalized precedence graph, containing only relations

**while** $\exists r \in R : r$ is a compound relation {

- replace $r$ by the sequence of relations it represents

} 

**return** $R$

- unpacks the compound relations
- required to get a real join tree as final result
Sample Algorithm Execution

Input: query graph

the precedence graph includes the ranks

Step 1: precedence graph for $R_1$
Sample Algorithm Execution (2)

Step 1: precedence graph for $R_1$

$R_2$  $R_3$  $R_4$  $R_5$  $R_6$  $R_7$

\[
\begin{array}{c}
\frac{49}{50} \\
\frac{24}{25} \\
\frac{5}{6} \\
\frac{19}{20} \\
\frac{4}{5} \\
\frac{1}{2}
\end{array}
\]

Step 2: normalization

$R_1$  $R_2$  $R_3$  $R_4$  $R_5$  $R_{6,7}$

\[
\begin{array}{c}
\frac{49}{50} \\
\frac{24}{25} \\
\frac{5}{6} \\
\frac{9}{15}
\end{array}
\]

$rank(R_6) > rank(R_7)$, but $R_6 \rightarrow R_7$
Sample Algorithm Execution (3)

Step 2: normalization

\[ rank(R_6,7) < rank(R_5) \]
Sample Algorithm Execution (3)

Step 3: merging subchains

Step 4: normalization

\[ rank(R_4) > rank(R_{6,7}), \text{ but } R_4 \rightarrow R_{6,7} \]
Sample Algorithm Execution (4)

Step 4: normalization  Step 5: merging subchains

\[\text{rank}(R_{4,6,7}) < \text{rank}(R_5) < \text{rank}(R_3) < \text{rank}(R_2)\]
Sample Algorithm Execution (5)

Step 5: merging subchains  Step 6: denormalization

Algorithm has to continue for all other root relations.
Maximum Value Precedence Algorithm

- greedy heuristics can produce poor results
- IKKBZ only support acyclic queries and ASI cost functions
- Maximum Value Precedence (MVP) [4] algorithm is a polynomial time heuristic with good results
- considers join ordering a graph theoretic problem
Directed Join Graph

Given a conjunctive query with predicates $P$.

- for all join predicates $p \in P$, we denote by $R(p)$ the relations whose attributes are mentioned in $p$.
- the \textit{directed join graph} of the query is a triple $G = (V, E_p, E_v)$, where $V$ is the set of predicates and $E_p$ and $E_v$ are sets of directed edges defined as follows
  - for any nodes $u, v \in V$, if $R(u) \cap R(v) \neq \emptyset$ then $(u, v) \in E_p$ and $(v, u) \in E_p$
  - if $R(u) \cap R(v) = \emptyset$ then $(u, v) \in E_v$ and $(v, u) \in E_v$
  - edges in $E_p$ are called \textit{physical edges}, those in $E_v$ \textit{virtual edges}

Note: all nodes $u, v$ there is an edge $(u, v)$ that is either physical or virtual. Hence, $G$ is a clique.
Examples: Spanning Tree and Join Tree

- every spanning tree in the directed join graph leads to a join tree
Examples: Spanning Tree and Join Tree (2)
Examples: Spanning Tree and Join Tree (3)

- spanning tree does not correspond to a (effective) join tree!
Effective Spanning Trees

It can be shown that a spanning tree $T = (V, E)$ is effective, it is satisfies the following conditions:

1. $T$ is a binary tree
2. for all inner nodes $v$ and nodes $u$ with $(u, v) \in E$:
   $\mathcal{R}(T(u)) \cap \mathcal{R}(v) \neq \emptyset$
3. for all nodes $v, u_1, u_2$ with $u_1 \neq u_2, (u_1, v) \in E$ and $(u_2, v) \in E$ one of the following conditions holds:
   3.1 $((\mathcal{R}(T(u_1)) \cap \mathcal{R}(v)) \cap (\mathcal{R}(T(u_2)) \cap \mathcal{R}(v))) = \emptyset$ or
   3.2 $(\mathcal{R}(T(u_1)) = \mathcal{R}(v)) \vee (\mathcal{R}(T(u_2)) = \mathcal{R}(v))$

We denote by $T(v)$ the partial tree rooted at $v$. 
Adding Weights to the Edges

For two nodes $v, u \in V$ we define $u \sqcap v = R(u) \cap R(v)$

- for simplicity, we assume that every predicate involves exactly two relations
- then for all $u, v \in V$, $a \sqcap v$ contains a single relation (or none)

Let $v \in V$ be a node with $R(v) = \{R_i, R_j\}$

- we abbreviate $R_i \Join_v R_j$ by $\Join_v$

Using these notations, we can attach weights to the edges to define the *weighted directed join graph*. 
Adding Weights to the Edges (2)

Let $G = (V, E_p, E_v)$ be a directed join graph for a conjunctive query with join predicates $P$. The weighted directed join graph is derived from $G$ by attaching a weight to each edge as follows:

- Let $(u, v) \in E_p$ be a physical edge. The weight $w_{u,v}$ of $(u, v)$ is defined as
  $$ w_{u,v} = \frac{|\Join u|}{|u \cap v|} $$

- For virtual edges $(u, v) \in E_v$, we define
  $$ w_{u,v} = 1 $$

Note that $w_{u,v}$ is not symmetric.
**Remark on Edge Weights**

The weights of physical edges are equal to the $s_i$ used in the IKKBZ-Algorithm.

Assume $R(u) = \{R_1, R_2\}$, $R(v) = \{R_2, R_3\}$. Then

$$w_{u,v} = \frac{|\bigtriangleup_u|}{|u \cap v|} = \frac{|R_1 \bigtriangleup R_2|}{|R_2|} = \frac{f_{1,2}|R_1||R_2|}{|R_2|} = f_{1,2}|R_1|$$

Hence, if the join $R_1 \bigtriangleup_u R_2$ is executed before the join $R_2 \bigtriangleup_v R_3$, the input size to the latter join changes by a factor of $w_{u,v}$. 

Adding Weights to the Nodes

- the weight of a node reflects the change in cardinality to be expected when certain other joins have been executed before
- it depends on a (partial) spanning tree $S$

Given $S$, we denote by $\Join^S_{p_{i,j}}$ the result of the join $\Join_{p_{i,j}}$ if all joins preceding $p_{i,j}$ in $S$ have been executed. Then the weight attached to node $p_{i,j}$ is defined as

$$w(p_{i,j}, S) = \frac{|\Join^S_{p_{i,j}}|}{|R_i \Join_{p_{i,j}} R_j|}$$

For empty sequences we define $w(p_{i,j}, \epsilon) = |R_i \Join_{p_{i,j}} R_j|$. Similarly, we define the cost of a node $p_{i,j}$ depending on other joins preceding it in some given spanning tree $S$. We denote this by $C(p_{i,j}, S)$.

- the actual cost function can be chosen arbitrarily
- if we have several join implementations: take the minimum
Algorithm Overview

The algorithm builds an effective spanning tree in two phases:

1. it takes those edges with a weight $< 1$
2. it adds the remaining edges

keeping track of effectiveness during the process.

- rational: weight $< 1$ is good
- decreases the work for later operators
- should be done early
- increasing intermediate results as late as possible
MVP Algorithm

MVP($G$)

**Input:** a weighted directed join graph $G = (V, E_p, E_v)$

**Output:** an effective spanning tree

$Q_1$ = a priority queue for nodes, largest $w$ first

$Q_2$ = a priority queue for nodes, smallest $w$ first

insert all nodes in $V$ to $Q_1$

$G' = (V', E')$ with $V' = V$ and $E' = E_p$ // working graph

$S = (V_S, E_s)$ with $V_s = V$ and $E_s = \emptyset$ // result

MVP-Phase1($G, G', S, Q_1, Q_2$)

MVP-Phase2($G, G', S, Q_1, Q_2$)

return \( S \)
MVP Algorithm (2)

MVP-Phase1($G, G', S, Q_1, Q_2$)

**Input:** state from MVP

**Output:** modifies the state

while $|Q_1| > 0 \land |E_s| < |V| - 1$ {

$v = \text{head of } Q_1$

$U = \{ u | (u, v) \in E' \land w_{u,v} < 1 \land (V, E_S \cup \{(u, v)\}) \text{ is acyclic and effective} \}$

if $U = \emptyset$ {

$Q_1 = Q_1 \setminus \{v\}$

$Q_2 = Q_2 \cup \{v\}$

} else {

$u = \text{arg max}_{u \in U} C(\nabla_v, S) - C(\nabla_v, (V, E_S \cup \{(u, v)\}))$

MVPUpdate($G, G', S, (u, v)$)

recompute $w$ for $v$ and its ancestors

}
MVP Algorithm (3)

MVP-Phase2($G, G', S, Q_1, Q_2$)

**Input:** state from MVP

**Output:** modifies the state

**while** $|Q_2| > 0 \land |E_s| < |V| - 1$ {

$v = \text{head of } Q_2$

$U = \{(x, y) | (x, y) \in E' \land (x = v \lor y = v) \land (V, E_S \cup \{(x, y)\}) \text{ is acyclic and effective}\}$

$(x, y) = \arg\min_{(x, y) \in U} C(x, y, (V, E_S \cup \{(x, y)\})) - C(x, y, S)$

MVPUpdate($G, G', S, (x, y)$)

recompute $w$ for $y$ and its ancestors

}
MVP Algorithm (4)

MVPUpdate($G, G', S, (u, v)$)

**Input:** state from MVP, an edge to be added to $S$

**Output:** modifies the state

$E_S = E_S \cup \{(u, v)\}$

$E' = E' \setminus \{(u, v), (v, u)\}$

$E' = E' \setminus \{(u, w)| (u, w) \in E'\}$

$E' = E' \cup \{(v, w)| (u, w) \in E_p, (v, w) \in E_v\}$

**if** $v$ has two incoming edges in $S$ **{**

$E' = E' \setminus \{(w, v)| (w, v) \in E'\}$

**}**

**if** $v$ has one outflowing edge in $S$ **{**

$E' = E' \setminus \{(v, w)| (v, w) \in E'\}$

**}**

- checks that $S$ is a tree (one parent, at most two children)
- detects transitive physical edges
Dynamic Programming

Basic premise:
- optimality principle
- avoid duplicate work

A very generic class of approaches:
- all cost functions (as long as optimality principle holds)
- left-deep/bushy, with/without cross products
- finds the optimal solution

Concrete algorithms can be more specialized of course.
Optimality Principle

Consider the two joins trees

$$(((R_1 \bowtie R_2) \bowtie R_3) \bowtie R_4) \bowtie R_5$$

and

$$(((R_3 \bowtie R_1) \bowtie R_2) \bowtie R_4) \bowtie R_5$$

- if we know that $$((R_1 \bowtie R_2) \bowtie R_3)$$ is cheaper than $$((R_3 \bowtie R_1) \bowtie R_2)$$, we know that the first join is cheaper than the second join.
- hence, we could avoid generating the second alternative and still won’t miss the optimal join tree.
Optimality Principle (2)

More formally, the optimality for join ordering:

Let $T$ be an optimal join tree for relations $R_1, \ldots, R_n$. Then, every subtree $S$ of $T$ must be an optimal join tree for the relations contained in it.

- optimal substructure: the optimal solution for a problem can be constructed from optimal solutions to its subproblems
- not true with physical properties (but can be fixed)
Overview Dynamic Programming Strategy

• generate optimal join trees bottom up
• start from optimal join trees of size one (relations)
• build larger join trees by (re-)using those of smaller sizes

To keep the algorithms concise, we use a subroutine \textit{CreateJoinTree} that joins two trees.
Creating Join Trees

CreateJoinTree( \( T_1, T_2 \) )

**Input:** two (optimal) join trees \( T_1, T_2 \)
for linear trees: assume that \( T_2 \) is a single relation

**Output:** an (optimal) join tree for \( T_1 \bowtie T_2 \)

\( B = \emptyset \)

for each \( impl \in \{ \) applicable join implementations \( \} \) {
  if \( \neg \)right-deep only {
    \( B = B \cup \{ T_1 \bowtie^{impl} T_2 \} \)
  }

  if \( \neg \)left-deep only {
    \( B = B \cup \{ T_2 \bowtie^{impl} T_1 \} \)
  }
}

return \( \arg \min_{T \in B} C(T) \)
Search Space with Sharing under Optimality Principle

\{R_1, R_2, R_3, R_4\}

\{R_1, R_2, R_4\}
\{R_1, R_2, R_3\}
\{R_1, R_3, R_4\}
\{R_1, R_2, R_4\}
\{R_1, R_3, R_4\}
\{R_2, R_3, R_4\}
\{R_1, R_3, R_4\}
\{R_1, R_2, R_3\}
\{R_1, R_2, R_4\}
\{R_1, R_2, R_3, R_4\}
Generating Linear Trees

- a (left-deep) linear tree $T$ with $|T| > 1$ has the form $T' \Join R_i$, with $|T| = |T'| + 1$
- if $T$ is optimal, $T'$ must be optimal too
- basic strategy: find the optimal $T$ by joining all optimal $T'$ with $T \setminus T'$

enumeration order varies between algorithms
Generating Linear Trees (2)

DPsizeLinear(R)

Input: a set of relations $R = \{R_1, \ldots, R_n\}$ to be joined

Output: an optimal left-deep (right-deep, zig-zag) join tree

$B = \text{an empty DP table } 2^R \rightarrow \text{join tree}$

for each $R_i \in R$

$$B[\{R_i\}] = R_i$$

for each $1 < s \leq n$ ascending {

for each $S \subset R$, $R_i \in R : |S| = s - 1 \land R_i \notin S$ {

if $\neg$cross products $\land \neg S$ connected to $R_i$ continue

$p_1 = B[S], p_2 = B[\{R_i\}]$

if $p_1 = \epsilon$ continue

$P = \text{CreateJoinTree}(p_1, p_2)$;

if $B[S \cup \{R_i\}] = \epsilon \lor C(B[S \cup \{R_i\}]) > C(P)$

$B[S \cup \{R_i\}] = P$

}

}$}

return $B[\{R_1, \ldots, R_n\}]$
Order in which Subtrees are generated

The ordering in which subtrees are generated does not matter as long as the following condition is not violated:

\[
\text{Let } S \text{ be a subset of } \{R_1, \ldots, R_n\}. \text{ Then, before a join tree for } S \text{ can be generated, the join trees for all relevant subsets of } S \text{ must already be available.}
\]

- \textit{relevant} means that they are valid subproblems by the algorithm
- usually this means connected (no cross products)
Generation in Integer Order

- can be done very efficiently
- set representation is just a number

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>000</td>
<td>{}</td>
</tr>
<tr>
<td>001</td>
<td>{R_1}</td>
</tr>
<tr>
<td>010</td>
<td>{R_2}</td>
</tr>
<tr>
<td>011</td>
<td>{R_1, R_2}</td>
</tr>
<tr>
<td>100</td>
<td>{R_3}</td>
</tr>
<tr>
<td>101</td>
<td>{R_1, R_3}</td>
</tr>
<tr>
<td>110</td>
<td>{R_2, R_3}</td>
</tr>
<tr>
<td>111</td>
<td>{R_1, R_2, R_3}</td>
</tr>
</tbody>
</table>
Generating Linear Trees (3)

\text{DPsubLinear}(R)

\textbf{Input:} a set of relations \( R = \{R_1, \ldots, R_n\} \) to be joined

\textbf{Output:} an optimal left-deep (right-deep, zig-zag) join tree

\( B = \) an empty DP table \( 2^R \rightarrow \) join tree

\textbf{for each} \( R_i \in R \)

\( B[\{R_i\}] = R_i \)

\textbf{for each} \( 1 < i \leq 2^n - 1 \) ascending \{ 

\( S = \{R_j \in R | (\lfloor i/2^{j-1} \rfloor \mod 2) = 1 \} \)

\textbf{for each} \( R_j \in S \) 

\textbf{if} \( \neg \text{cross products} \land \neg S \setminus \{R_j\} \) connected to \( R_j \) \textbf{continue}

\( p_1 = B[S \setminus \{R_j\}], p_2 = B[\{R_j\}] \)

\textbf{if} \( p_1 = \epsilon \) \textbf{continue}

\( P = \text{CreateJoinTree}(p_1, p_2); \)

\textbf{if} \( B[S] = \epsilon \lor C(B[S]) > C(P) \) \( B[S] = P \)

\}

\}

\textbf{return} \( B[\{R_1, \ldots, R_n\}] \)
Generating Bushy Trees

- a bushy tree $T$ with $|T| > 1$ has the form $T_1 \Join T_2$, with $|T| = |T_1| + |T_2|$
- if $T$ is optimal, both $T_1$ and $T_2$ must be optimal too
- basic strategy: find the optimal $T$ by joining all pairs of optimal $T_1$ and $T_2$
Generating Bushy Trees (2)

DPsize($R$)

**Input:** a set of relations $R = \{ R_1, \ldots, R_n \}$ to be joined

**Output:** an optimal bushy join tree

$B$ = an empty DP table $2^R \rightarrow$ join tree

**for each** $R_i \in R$

$B[\{R_i\}] = R_i$

**for each** $1 < s \leq n$ ascending 

**for each** $S_1, S_2 \subset R : |S_1| + |S_2| = s$

 - if ($\neg$cross products $\land \neg S_1$ connected to $S_2$) $\lor$ ($S_1 \cap S_2 \neq \emptyset$) continue

  $p_1 = B[S_1], p_2 = B[S_2]$

  if $p_1 = \epsilon \lor p_2 = \epsilon$ continue

  $P = \text{CreateJoinTree}(p_1, p_2)$;

  if $B[S_1 \cup S_2] = \epsilon \lor C(B[S_1 \cup S_2]) > C(P)$

  $B[S_1 \cup S_2] = P$

**}**

**}**

**return** $B[\{R_1, \ldots, R_n\}]$
Generating Bushy Trees (3)

DPsub($R$)

**Input:** a set of relations $R = \{R_1, \ldots, R_n\}$ to be joined

**Output:** an optimal bushy join tree

$B =$ an empty DP table $2^R \rightarrow \text{join tree}$

**for each** $R_i \in R$

\[ B[\{R_i\}] = R_i \]

**for each** $1 < i \leq 2^n - 1$ ascending

\[ S = \{R_j \in R | ([i/2^{j-1}] \mod 2) = 1\} \]

**for each** $S_1 \subset S, S_2 = S \setminus S_1$

- if \( \neg \text{cross products} \land \neg S_1 \text{ connected to } S_2 \) continue
- $p_1 = B[S_1], p_2 = B[S_2]$
- if $p_1 = \epsilon \lor p_2 = \epsilon$ continue
- $P = \text{CreateJoinTree}(p_1, p_2)$
- if $B[S] = \epsilon \lor C(B[S]) > C(P) \quad B[S] = P$

**return** $B[\{R_1, \ldots, R_n\}]$
Efficient Subset Generation

If we use integers as set representation, we can enumerate all subsets of $S$ as follows:

$$S_1 = S \& (-S)$$

**do**

- $S_2 = S - S_1$
- // Do something with $S_1$ and $S_2$
- $S_1 = S \& (S_1 - S)$

**while** ($S_1 \neq S$)

- enumerates all subsets except $\emptyset$ and $S$ itself
- very fast
Remarks

- \text{DPsize/DPsizeLinear} does not really test for \( p_1 = \epsilon \)
- it keeps a list of plans for a given size
- candidates can be found very fast
- ensures polynomial time in some cases (we will look at it again)
- \text{DPsub/DPsubLinear} is faster if the problem is not polynomial, though
Memoization

- top-down formulation of dynamic programming
- recursive generation of join trees
- memoize already generated join trees to avoid duplicate work
- easier code
- sometimes more efficient (more knowledge, allows for pruning)
- but usually slower than dynamic programming
Memoization (2)

Memoization($R$)

**Input:** a set of relations $R = \{R_1, \ldots, R_n\}$ to be joined

**Output:** an optimal bushy join tree

$B$ = an empty DP table $2^R \rightarrow$ join tree

**for each** $R_i \in R$

$B[\{R_i\}] = R_i$

MemoizationRec($B, R$)

**return** $B[\{R_1, \ldots, R_n\}]$

- initializes the DP table and triggers the recursive search
- main work done during recursion
Memoization (3)

MemoizationRec($B, S$)

**Input:** a DP table $B$ and a set of relations $S$ to be joined

**Output:** an optimal bushy join tree for the subproblem

if $B[S] = \epsilon$ {
  for each $S_1 \subset S$, $S_2 = S \setminus S_1$
    $p_1 = \text{MemoizationRec}(B, S_1)$, $p_2 = \text{MemoizationRec}(B, S_2)$
    $P = \text{CreateJoinTree}(p_1, p_2)$
    if $B[S] = \epsilon \lor C(B[S]) > C(P)$ $B[S] = P$
  }
return $B[S]$

- checks for connectedness omitted
Dynamic Programming - Connected Subgraphs

- DP a very versatile strategy
- common usage scenario: bushy, no cross products
- DPsize and DPsub support it, of course, but not optimal
- enumeration order does not consider the query graph
- many pairs have to be pruned due to connectedness
- especially bad for DPsub

Solution: consider the query graph structure during DP enumeration [5]
Asymptotic Search Space

DPsize:

- organize DP by the size of the join tree
- problem: only few DP slots, many pairs considered

good algorithm for chains, very bad for cliques:

<table>
<thead>
<tr>
<th></th>
<th>chains</th>
<th>cycles</th>
<th>stars</th>
<th>cliques</th>
</tr>
</thead>
<tbody>
<tr>
<td>pairs</td>
<td>$O(n^4)$</td>
<td>$O(n^4)$</td>
<td>$O(4^n)$</td>
<td>$O(4^n)$</td>
</tr>
</tbody>
</table>

DPsub:

- organize DP by the set of relations involved
- problem: always $2^n$ DP slots, fixed enumeration

good algorithm for cliques, but adapts badly:

<table>
<thead>
<tr>
<th></th>
<th>chains</th>
<th>cycles</th>
<th>stars</th>
<th>cliques</th>
</tr>
</thead>
<tbody>
<tr>
<td>pairs</td>
<td>$O(2^n)$</td>
<td>$O(n2^n)$</td>
<td>$O(3^n)$</td>
<td>$O(3^n)$</td>
</tr>
</tbody>
</table>
Observation

DPsize and DPsub generate many pairs that are pruned anyway (connectedness, overlap).

Typical pruned pairs (chain with 4 relations):

- not connected
- not disjoint
- invalid subproblems

Last example ⇒ every join partner must be a connected subgraph:
Graph Theoretic Approach

- reformulation as graph theoretic problem:
- enumerate all connected subgraphs of the query graph
- for each subgraph enumerate all other connected subgraphs that are disjoint but connected to it
- each connected subgraph - complement pair (ccp) can be joined
- enumerate them suitable for DP ⇒ DP algorithm

Algorithm adapts naturally to the graph structure:

<table>
<thead>
<tr>
<th>pairs</th>
<th>chains</th>
<th>cycles</th>
<th>stars</th>
<th>cliques</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O(n^3)$</td>
<td>$O(n^3)$</td>
<td>$O(n^{2n})$</td>
<td>$O(3^n)$</td>
<td></td>
</tr>
</tbody>
</table>

Lohman et al: \#ccp is a lower bound for all DP enumeration algorithms
DP Algorithm using Connected Subgraphs

If we can efficiently enumerate all connected subgraphs/connected complement pairs, the resulting DP algorithm is:

DPccp(R)

**Input:** a connected query graph with relations $R = \{R_0, \ldots, R_{n-1}\}$

**Output:** an optimal bushy join tree

$B = \text{an empty DP table } 2^R \rightarrow \text{join tree}$

**for** $\forall R_i \in R$

$B[\{R_i\}] = R_i$

**for** $\forall$ csg-cmp-pairs $(S_1, S_2)$, $S = S_1 \cup S_2$ {

$p_1 = B[S_1], p_2 = B[S_2]$

$P = \text{CreateJoinTree}(p_1, p_2)$;

if $B[S] = \epsilon \lor C(B[S]) > C(P)$

$B[S] = P$

}

**return** $B[\{R_0, \ldots, R_{n-1}\}]$

The main problem is enumerating the pairs.
## Effect on Search Space

Absolute number of generated pairs

<table>
<thead>
<tr>
<th>$n$</th>
<th>$n$</th>
<th>Chain</th>
<th>Star</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DPccp</td>
<td>DPsub</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>20</td>
<td>84</td>
</tr>
<tr>
<td>10</td>
<td>165</td>
<td>165</td>
<td>3,962</td>
</tr>
<tr>
<td>15</td>
<td>560</td>
<td>560</td>
<td>130,798</td>
</tr>
<tr>
<td>20</td>
<td>1,330</td>
<td>1,330</td>
<td>4,193,840</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Clique</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>Chain</td>
</tr>
<tr>
<td></td>
<td>DPccp</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>10</td>
<td>405</td>
</tr>
<tr>
<td>15</td>
<td>1,470</td>
</tr>
<tr>
<td>20</td>
<td>3,610</td>
</tr>
</tbody>
</table>
Enumerating Connected Subgraphs

- two steps: enumerate all connected subgraphs, enumerate disjoint but connected subgraphs for a given one ⇒ pairs
- enumerate all pairs, enumerate no duplicates, enumerate for DP
- if \((a, b)\) is enumerated, do not enumerate \((b, a)\)
- requires total ordering of connected subgraphs
- preparation: label nodes breadth-first from 0 to \(n - 1\)

Preliminaries, given query graph \(G = (V, E)\):

\[
V = \{v_0, \ldots, v_{n-1}\}
\]

\[
\mathcal{N}(V') = \{v' \mid v \in V' \land (v, v') \in E\}
\]

\[
\mathcal{B}_i = \{v_j \mid j \leq i\}
\]
Enumerating Connected Subgraphs (2)

EnumerateCsg(\(G\))

\[
\text{for all } i \in [n-1, \ldots, 0] \text{ descending } \{
\begin{align*}
&\text{emit } \{v_i\}; \\
&\text{EnumerateCsgRec} (G, \{v_i\}, \mathcal{B}_i);
\end{align*}
\}
\]

EnumerateCsgRec(\(G, S, X\))

\[
N = \mathcal{N}(S) \setminus X;
\]

\[
\text{for all } S' \subseteq N, S' \neq \emptyset, \text{ enumerate subsets first } \{
\begin{align*}
&\text{emit } (S \cup S'); \\
&\text{EnumerateCsgRec} (G, (S \cup S'), (X \cup N));
\end{align*}
\}
\]
Enumerating Connected Subgraphs (2)

EnumerateCsg\(G\)
for all \(i \in [n - 1, \ldots, 0]\) descending \{ 
  \text{emit} \ \{v_i\}; 
  \text{EnumerateCsgRec}(G, \{v_i\}, B_i); 
\}

Choose all nodes as enumeration start node once

EnumerateCsgRec\(G, S, X\)
\(N = \mathcal{N}(S) \setminus X;\)
for all \(S' \subseteq N, S' \neq \emptyset\), enumerate subsets first \{ 
  \text{emit} \ (S \cup S'); 
\}
for all \(S' \subseteq N, S' \neq \emptyset\), enumerate subsets first \{ 
  \text{EnumerateCsgRec}(G, (S \cup S'), (X \cup N)); 
\}
Enumerating Connected Subgraphs (2)

EnumerateCsg($G$)

for all $i \in [n-1, \ldots, 0]$ descending {
    emit $\{v_i\}$;
    EnumerateCsgRec($G$, $\{v_i\}$, $B_i$);
}

EnumerateCsgRec($G$, $S$, $X$)

$N = \mathcal{N}(S) \setminus X$;

for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    emit $(S \cup S')$;
}

for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    EnumerateCsgRec($G$, $(S \cup S')$, $(X \cup N)$);
}
Enumerating Connected Subgraphs (2)

EnumerateCsg($G$)
for all $i \in [n - 1, \ldots, 0]$ descending {
    emit $\{v_i\}$;
    EnumerateCsgRec($G$, $\{v_i\}$, $\mathcal{B}_i$);
}

EnumerateCsgRec($G$, $S$, $X$)
$N = \mathcal{N}(S) \setminus X$; 
for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    emit $(S \cup S')$;
}
for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    EnumerateCsgRec($G$, $(S \cup S')$, $(X \cup N)$);
}
Enumerating Connected Subgraphs (2)

EnumerateCsg\( (G) \)
for all \( i \in [n-1, \ldots, 0] \) descending \{
  emit \{ v_i \} ;
  EnumerateCsgRec\( (G, \{ v_i \}, B_i) \);
\}

EnumerateCsgRec\( (G, S, X) \)
\( N = \mathcal{N}(S) \setminus X ; \)
for all \( S' \subseteq N, S' \neq \emptyset \), enumerate subsets first \{
  emit \( (S \cup S') \) ;
\}
for all \( S' \subseteq N, S' \neq \emptyset \), enumerate subsets first \{
  EnumerateCsgRec\( (G, (S \cup S'), (X \cup N)) \) ;
\}

Prohibit nodes with smaller labels. Thus the set of valid nodes increases over time
Enumerating Connected Subgraphs (2)

\text{EnumerateCsg}(G)
\text{for all } i \in [n - 1, \ldots, 0] \text{ descending } \{
    \text{emit } \{v_i\}; \\
    \text{EnumerateCsgRec}(G, \{v_i\}, \mathcal{B}_i);
\}

\text{EnumerateCsgRec}(G, S, X)
N = \mathcal{N}(S) \setminus X;
\text{for all } S' \subseteq N, S' \neq \emptyset, \text{ enumerate subsets first } \{
    \text{emit } (S \cup S');
\}
\text{for all } S' \subseteq N, S' \neq \emptyset, \text{ enumerate subsets first } \{
    \text{EnumerateCsgRec}(G, (S \cup S'), (X \cup N));
\}
Enumerating Connected Subgraphs (2)

EnumerateCsg\( (G) \)

\textbf{for all} \( i \in [n-1, \ldots, 0] \) \textbf{descending} \{ \\
\text{emit} \{v_i\}; \\
\text{EnumerateCsgRec}(G, \{v_i\}, \mathcal{B}_i); \\
\}

EnumerateCsgRec\( (G, S, X) \)

\( N = \mathcal{N}(S) \setminus X; \)

\textbf{for all} \( S' \subseteq N, S' \neq \emptyset \), enumerate subsets first \{ \\
\text{emit} (S \cup S'); \\
\}

\textbf{for all} \( S' \subseteq N, S' \neq \emptyset \), enumerate subsets first \{ \\
\text{EnumerateCsgRec}(G, (S \cup S'), (X \cup N)); \\
\}
Enumerating Connected Subgraphs (2)

EnumerateCsg(G)
for all $i \in [n-1, \ldots, 0]$ descending {
    emit $\{v_i\}$;
    EnumerateCsgRec(G, $\{v_i\}$, $B_i$);
}

EnumerateCsgRec(G, S, X)
$N = \mathcal{N}(S) \setminus X$;
for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    emit $(S \cup S')$;
}
for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    EnumerateCsgRec(G, $(S \cup S')$, $(X \cup N)$);
}
Enumerating Connected Subgraphs (2)

EnumerateCsg\((G)\)

for all \(i \in [n - 1, \ldots, 0]\) descending {
    emit \(\{v_i\}\);
    EnumerateCsgRec\((G, \{v_i\}, B_i)\);
}

EnumerateCsgRec\((G, S, X)\)

\[N = \mathcal{N}(S) \setminus X;\]

for all \(S' \subseteq N, S' \neq \emptyset\), enumerate subsets first {
    emit \((S \cup S')\);
}

for all \(S' \subseteq N, S' \neq \emptyset\), enumerate subsets first {
    EnumerateCsgRec\((G, (S \cup S'), (X \cup N))\);
}
Enumerating Connected Subgraphs (2)

EnumerateCsg(G)

for all $i \in [n - 1, \ldots, 0]$ descending {
    emit $\{v_i\}$;
    EnumerateCsgRec(G, $\{v_i\}$, $\mathcal{B}_i$);
}

EnumerateCsgRec(G, S, X)

$N = \mathcal{N}(S) \setminus X$;

for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    emit $(S \cup S')$;
}

for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    EnumerateCsgRec(G, $(S \cup S')$, $(X \cup N)$);
}

Add all combinations to the subgraph and emit the new subgraph
Enumerating Connected Subgraphs (2)

\[
\text{EnumerateCsg}(G) \\
\text{for all } i \in [n-1, \ldots, 0] \text{ descending } \{ \\
\quad \text{emit } \{v_i\}; \\
\quad \text{EnumerateCsgRec}(G, \{v_i\}, B_i); \\
\} \\
\]

\[
\text{EnumerateCsgRec}(G, S, X) \\
N = \mathcal{N}(S) \setminus X; \\
\text{for all } S' \subseteq N, S' \neq \emptyset, \text{ enumerate subsets first } \{ \\
\quad \text{emit } (S \cup S'); \\
\} \\
\text{for all } S' \subseteq N, S' \neq \emptyset, \text{ enumerate subsets first } \{ \\
\quad \text{EnumerateCsgRec}(G, (S \cup S'), (X \cup N)); \\
\} \\
\]

Add all combinations to the subgraph and emit the new subgraph
Enumerating Connected Subgraphs (2)

EnumerateCsg(G)
for all $i \in [n-1, \ldots, 0]$ descending {
    emit \{v_i\};
    EnumerateCsgRec(G, \{v_i\}, B_i);
}

EnumerateCsgRec(G, S, X)
$N = \mathcal{N}(S) \setminus X$;
for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    emit (S ∪ S');
}
for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    EnumerateCsgRec(G, (S ∪ S'), (X ∪ N));
}

Add all combinations to the subgraph and emit the new subgraph
Enumerating Connected Subgraphs (2)

EnumerateCsg(G)
for all \( i \in [n - 1, \ldots, 0] \) descending {
    emit \( \{v_i\}\);
    EnumerateCsgRec(G, \{v_i\}, \mathcal{B}_i);
}

EnumerateCsgRec(G, S, X)
\( N = \mathcal{N}(S) \setminus X; \)
for all \( S' \subseteq N, S' \neq \emptyset \), enumerate subsets first {
    emit \( (S \cup S')\);
}
for all \( S' \subseteq N, S' \neq \emptyset \), enumerate subsets first {
    EnumerateCsgRec(G, \( (S \cup S')\), \( (X \cup N)\));
}
Enumerating Connected Subgraphs (2)

EnumerateCsg(G)

for all $i \in [n - 1, \ldots, 0]$ descending {
    emit $\{v_i\}$;
    EnumerateCsgRec(G, $\{v_i\}$, $B_i$);
}

EnumerateCsgRec(G, S, X)

$N = \mathcal{N}(S) \setminus X$;

for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    emit $(S \cup S')$;
}

for all $S' \subseteq N$, $S' \neq \emptyset$, enumerate subsets first {
    EnumerateCsgRec(G, $(S \cup S')$, $(X \cup N)$);
}
Enumerating Complementary Subgraphs

\text{EnumerateCmp}(G, S_1)
\begin{align*}
X &= B_{\min}(S_1) \cup S_1; \\
N &= N(S_1) \setminus X; \\
\text{for all } (v_i \in N \text{ by descending } i) \{ \\
\text{emit } \{v_i\}; \\
\text{EnumerateCsgRec}(G, \{v_i\}, X \cup (B_i \cap N)); \\
\}
\end{align*}
**Enumerating Complementary Subgraphs**

EnumerateCmp\(G,S_1\)

\[
X = B_{\text{min}}(S_1) \cup S_1;
\]

\[
N = \mathcal{N}(S_1) \setminus X;
\]

for all \(v_i \in N\) by descending \(i\) {

- **emit** \(\{v_i\}\);
- EnumerateCsgRec\(G, \{v_i\}, X \cup (B_i \cap N)\);
}

Prohibit all nodes that will be start nodes later on and the primary subgraph
Enumerating Complementary Subgraphs

\[ \text{EnumerateCmp}(G, S_1) \]
\[ X = B_{\text{min}}(S_1) \cup S_1; \]
\[ N = \mathcal{N}(S_1) \setminus X; \]

for all \( (v_i \in N \text{ by descending } i) \) {
    \[ \text{emit } \{v_i\}; \]
    \[ \text{EnumerateCsgRec}(G, \{v_i\}, X \cup (B_i \cap N)); \]
}

Find all neighboring nodes that are not prohibited
Enumerating Complementary Subgraphs

EnumerateCmp\((G, S_1)\)

\[ X = B_{\text{min}}(S_1) \cup S_1; \]

\[ N = \mathcal{N}(S_1) \setminus X; \]

for all \((v_i \in N \text{ by descending } i)\) {
    \begin{align*}
    & \text{emit } \{ v_i \}; \\
    & \text{EnumerateCsgRec}(G, \{ v_i \}, X \cup (B_i \cap N));
    \end{align*}
}

Consider each of the nodes

\[ R_0 \]

\[ R_1 \]

\[ R_2 \]

\[ R_3 \]

\[ R_4 \]
Enumerating Complementary Subgraphs

$$\text{EnumerateCmp}(G, S_1)$$

$$X = B_{\min}(S_1) \cup S_1;$$

$$N = N(S_1) \setminus X;$$

for all $v_i \in N$ by descending $i$) {

    emit \{v_i\};

    EnumerateCsgRec($G$, \{v_i\}, $X \cup (B_i \cap N)$);

}
Enumerating Complementary Subgraphs

EnumerateCmp(G, S₁)

\[ X = B_{\text{min}}(S₁) \cup S₁; \]
\[ N = \mathcal{N}(S₁) \setminus X; \]

for all \( v_i \in N \) by descending \( i \) {
    emit \( \{v_i\} \);
    EnumerateCsgRec(G, \{v_i\}, X \cup (B_i \cap N));
}

Recursively increase the subgraph re-using EnumerateCsgRec
Enumerating Complementary Subgraphs

\[ \text{EnumerateCmp}(G, S_1) \]
\[ X = B_{\min}(S_1) \cup S_1; \]
\[ N = N(S_1) \setminus X; \]
\[ \text{for all } (v_i \in N \text{ by descending } i) \{ \]
\[ \quad \text{emit } \{v_i\}; \]
\[ \quad \text{EnumerateCsgRec}(G, \{v_i\}, X \cup (B_i \cap N)); \]
\[ \} \]

Again prohibit nodes with a smaller label to prevent duplicates.
Enumerating Complementary Subgraphs

EnumerateCmp($G, S_1$)

$X = B_{\text{min}}(S_1) \cup S_1$;

$N = \mathcal{N}(S_1) \setminus X$;

for all ($v_i \in N$ by descending $i$) {
    emit $\{v_i\}$;
    EnumerateCsgRec($G$, $\{v_i\}$, $X \cup (B_i \cap N)$);
}

- EnumerateCsg + EnumerateCmp produce all ccp
- resulting algorithm DPccp considers exactly $\#ccp$ pairs
- which is the lower bound for all DP enumeration algorithms
Remarks

- DPsize is good for chains, DPsub for cliques
- implementation of DPccp is more involved
- each enumeration step must be fast (ideally $O(1)$, at most $O(n)$, where $n$ is the number of relations)
- but benefits are huge
- DPccg adopts to query graph structure
- considers minimal number of pairs
- especially for "in-between queries" (e.g. stars) much faster
Beyond (Regular) Query Graphs

Some queries are more complex

```
select *
from  R1 r1, R2 r2, R3 r3,
      R4 r4, R5 r5, R6 r6
where r1.a = r2.a and r2.b = r3.c and
     r4.d = r5.d and r5.e = r6.e and
     abs(r1.f + r3.f) = abs(r4.g + r6.g)
```

- does not induce a graph but a hyper-graph
- graph based DP algorithm not directly applicable
- generic DP algorithms work, but not as efficient
Handling Hypergraphs

A hypergraph is a pair $H = (V, E)$ such that

1. $V$ is a non-empty set of nodes and
2. $E$ is a set of hyperedges, where a hyperedge is an unordered pair $(u, v)$ of non-empty subsets of $V$ ($u \subset V$ and $v \subset V$) with the additional condition that $u \cap v = \emptyset$.

Nodes in $V$ are totally ordered via an (arbitrary) relation $\prec$.

- enumeration is performed by decreasing $\prec$
- $\prec$ orders the search space (DP order, duplicates)
Handling Hypergraphs (2)

In principle same approach as for regular graphs:

- start with one node
- expand recursively by following edges

Problem:
- hyperedges are n:m edges
- where to expand to from \{R_1, R_2, R_3\}?
- must still guarantee DP order
Handling Hypergraphs - Neighborhood

When computing the neighborhood, choose representatives:

- a hyperedge "leads" to the least node (regarding $\prec$)
- therefore $N(\{ R_1, R_2, R_3 \}) = \{ R_4 \}$
- ensures DP order (and prevents duplicates)

But:

- leads to (temporarily) disconnected graphs
- $\{ R_1, R_2, R_3, R_4 \}$ is not connected
- must expand further until $R_6$ reached

Requires checks for connectedness

- can exploit the DP table for cheap tests
- if it is connected, a DP entry must exist
Non-Inner Joins

Some queries use non-inner joins:

- either explicitly (OUTER JOIN etc.) or implicitly (unnesting etc.)
- are not freely reorderable

Must be taken into account during join ordering.
Non-Inner Joins - Reordering Constraints

Examine pair-wise reorderings of operators

- for all $\circ_1, \circ_2$, check if $(R \circ_1 S) \circ_2 T \equiv R \circ_1 (S \circ_2 T)$
- assume syntax constraints are satisfied

Gives a big compatibility matrix

<table>
<thead>
<tr>
<th>$\bowtie$</th>
<th>$\bowtie$</th>
<th>$\bowtie$</th>
<th>$\triangleright$</th>
<th>$\bowtie$</th>
<th>$\bowtie$</th>
<th>$\bowtie$</th>
<th>...</th>
</tr>
</thead>
</table>
| $\bowtie$ | $+$       | $+$       | $-$              | $+$       | $+$       | $+$       | $+$ | ...
| $\bowtie$ | $-$       | $+$       | $-$              | $-$       | $-$       | $-$       | $-$ | ...
| $\bowtie$ | $-$       | $+$       | $+$              | $-$       | $-$       | $-$       | $-$ | ...
| $\triangleright$ | $-$       | $-$       | $-$              | $-$       | $-$       | $-$       | $-$ | ...
| $\bowtie$ | $-$       | $-$       | $-$              | $-$       | $-$       | $-$       | $-$ | ...
| $\bowtie$ | $-$       | $-$       | $-$              | $-$       | $-$       | $-$       | $-$ | ...
| $\bowtie$ | $-$       | $-$       | $-$              | $-$       | $-$       | $-$       | $-$ | ...
| ... | ... | ... | ... | ... | ... | ... | ... |
Non-Inner Joins - TESs

Extract reordering constraints from operator tree in two steps:

1. build the *syntactic eligibility set* (SES) for each operator
   ▶ set of relations that has to be in the input

```
+---+  +---+  +---+  +---+
| A |  | b=d |  | B |  | d=
|---|  |-----|  |---|  |---|
|---|  |     |  |---|  |     |
| a=c|  |     |  |---|  |     |
|---|  |     |  |---|  |     |
| a=b|  |     |  |---|  |     |
+---+  +---+  +---+  +---+
     |     |     |     |
     |     |     |     |
     +---+  +---+  +---+  +---+

SES

{B, D}

{A, C}

{A, B}
```
Non-Inner Joins - TESs

Extract reordering constraints from operator tree in two steps:

1. build the *syntactic eligibility set* (SES) for each operator
2. bottom up traversal, build the *total eligibility set* (TES)
   ▶ initialize TES with SES
   ▶ check for conflicts with other operators (can be in subtrees!)
   ▶ if conflict, add other TES to own TES

<table>
<thead>
<tr>
<th>SES</th>
<th>TES</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{ SES } )</td>
<td>( \text{ TES } )</td>
</tr>
<tr>
<td>{ B, D }</td>
<td>{ A, B, D }</td>
</tr>
<tr>
<td>{ A, C }</td>
<td>{ A, B, C }</td>
</tr>
<tr>
<td>{ A, B }</td>
<td>{ A, B }</td>
</tr>
</tbody>
</table>

TESs capture reordering restrictions by requiring relations, which imply operators.
Non-Inner Joins - Using TESs

Add the TES to the join edge

- operator ”requires” certain relations, so encode it like this
- constructs hyperedges (n:m)
- eliminates invalid reorderings from the search space

Original query graph from previous example: $C \leftarrow A \leftarrow B \leftarrow D$

After adding TESs to the edges: $C \leftarrow A \leftarrow B \lefthyphenated\rightarrow D$
Simplifying the Query Graph

The graph-based DP algorithm considers the minimal number of join-pairs
- we therefore cannot expect to get a better runtime for exact solutions
- many problems can be solved exactly, but not all
- depends on the structure of the query graph
- chains are simple, others, e.g., stars, are hard
- how to cope with these queries?

Greedy heuristics would work, but results are much worse than DP solutions.
Simplifying the Query Graph - General Idea

If the problem is too complex to solve exactly, simplify the query graph until it gets tractable.

- the query graph describes all join possibilities
- by modifying the query graph we can rule out some possibilities
- this reduces the search space and the optimization time
- we prefer modifications that are "safe"
- uses greedy steps only for the "easy" problems, then use DP

Note: "simplifying" means simpler for the optimizer.
For a human the query graph tends to get strange.
Simplifying a Star Query

graph

\[ R_0 \rightarrow R_1 \\ R_3 \ \ \ \ \ \ \ \ \ R_2 \]

joins

\[ R_0 \bowtie R_1 \]
\[ R_0 \bowtie R_2 \]
\[ R_0 \bowtie R_3 \]

original

search space size

6
Simplifying a Star Query

We decide to order $R_0 \bowtie R_1$ before $R_0 \bowtie R_2$ (introduces hyperedge)
Simplifying a Star Query

We decide to order $R_0 \Join R_1$ before $R_0 \Join R_3$ (introduces hyperedge)
Simplifying a Star Query

We decide to order \( \{ R_0, R_1 \} \bowtie R_2 \) before \( R_0 \bowtie R_3 \) (introduces hyperedge).
Performing A Simplification Step

Given a query graph \( G = (V, E) \)

1. examine all joins \( \Join_1, \Join_2 \in E \) that are neighboring
   - neighboring \( \approx \) have a relation in common (see [6])
2. make sure that \( \Join_2 \) could be ordered before \( \Join_1 \)
   - checks for contradictions, requires a fast cycle checker
3. compute the \( \text{orderingBenefit}(\Join_1, \Join_2) \)
   - this is the heuristical part, different benefit heuristics could be used
4. retain the \( S^L_1 \Join_1 S^R_1, S^L_2 \Join_2 S^R_2 \) with the maximal orderingBenefit
   - maintain priority queues to speed up repeated simplification
5. return \( G' = (V, E \setminus \{\Join_1\} \cup \{(S^L_1 \cup S^L_2 \cup S^R_2) \Join_1 S^R_1\}) \)

The resulting query graph is more restrictive, i.e., simpler.

(there are more cases due to different possible ways of neighboring)
Estimating the Ordering Benefit

We want to prefer orderings that are almost certainly a good idea. Therefore one approach is to maximize

$$\text{orderingBenefit}(X \Join_1 R_1, X \Join_2 R_2) = \frac{C((X \Join_1 R_1) \Join_2 R_2)}{C((X \Join_2 R_2) \Join_1 R_1)}$$

If we cannot compute $C$ due to missing information, use $C_{out}$. 
Adjusting the Problem Complexity

How much should we simplify?

- until optimization fits into resource constraints (memory or time)

How do we know when to stop simplifying?

- count the number of **connected subgraphs** of the query graph
  - directly determines memory, indirectly optimization time
  - stop counting when the limit is reached

Counting is fast, but not instantaneous

- counting 10,000 subgraphs in a query with 100 relations took \( \approx 5 \text{ms} \)
  - we cannot do this after every simplification

Exact limit depends on hardware, a reasonable choice is 10,000 connected subgraphs.
Full Optimization Algorithm

Given a Query Graph $G = (V, E)$ and a complexity budget $b$

1. compute a list $\bar{G}$ of query graphs
   - repeatedly call the simplification step, stop when no change

2. perform binary search over $\bar{G}$, find $G_b$
   - for the current element $G'$, $c = \#$connected subgraphs in $G'$ (count at most $b + 1$)
   - if $c > b$ increase, otherwise decrease

3. return $DPhyp(G_b)$

Simplifies as much as needed to meet the constraints, than uses DP.

(the algorithm does not materialize $\bar{G}$ explicitly, see [6])
Time/Quality Trade-off - Grid with 20 Relations

- as expected plan quality degrades at some point
- but optimization times drops off much earlier
Time/Quality Trade-off - Star with 20 Relations

- same optimization time behavior, but plan quality remains perfect
Adaptive Optimization using Search Space Linearization

- not all join problems are equal
- most queries are small, but we have a incredible long tail
- must handle all of them reasonably, with the correct expectations
- adapt the algorithm to the query complexity
Join ordering: Solved!

- *Dynamic Programming (DP)*, pioneered by Selinger et al. (1979)
- Large body of follow-up work
  - bushy plans
  - graph awareness
  - non-inner joins
  - top-down formulations
- *Exponential* runtime in general
- Only viable for relatively *small queries*
- Generated queries we are increasingly faced with tend to be *too large*
Solved?

- Huge search space (NP-Hard)
- Too hard to solve optimally
- Heuristics to the rescue!?

- Suddenly, if just slightly too large
- Likely to result in disastrous plans
- Not the end of the spectrum
Unsolved!

- **Tableau**: “Get Real: How Benchmarks Fail to Represent the Real World” (DBTEST 2018)
- Queries touching a few **hundred relations** are quite **common**
- **SAP**: 4,598 relations (BTW 2017)
Adaptive Optimization – The Big Picture

- For **performance** and **correctness** reasons: no cross products
Adaptive Optimization – The Big Picture

- **Query**
  - easy? yes: solve *optimally*
  - easy? no: medium?
    - medium? yes: novel technique: *search space linearization*
    - medium? no: gracefully introduce *greediness* to keep optimization time reasonable
Adaptive Optimization – The Big Picture

Query

- easy?
  - yes: solve *optimally*
  - no:
    - medium?
      - yes: novel technique: *search space linearization*
      - no: gracefully introduce *greediness* to keep optimization time reasonable
Adaptive Optimization – The Big Picture

- Query
- \( \leq 10K \) DP entries?
  - yes: solve *optimally* using DPhyp
  - no: medium?
    - no: cannot linearize
      - yes: novel technique: *search space linearization*
    - yes: gracefully introduce *greediness* to keep optimization time reasonable
Adaptive Optimization – The Big Picture

- **Query**
  - \( \leq 10K \text{ DP entries?} \)
  - no → **corner case?**
  - yes \( \leq 10K \text{ DP entries?} \)
  - no → **medium?**
  - yes → novel technique: search space linearization
  - no → gracefully introduce greediness to keep optimization time reasonable

- DPhyp
- GOO/DPhyp
Adaptive Optimization – The Big Picture

- Query

- ≤ 10K DP entries?
  - yes: solve \textit{optimally} using DPhyp
  - no: corner case?

- corner case?
  - yes: cannot linearize using GOO/DPhyp
  - no: ≤ 100 relations?

- ≤ 100 relations?
  - yes: novel technique: \textit{search space linearization}
    - linearized DP
  - no: gracefully introduce \textit{greediness} to keep optimization time reasonable
Adaptive Optimization – The Big Picture

- **Query**
  - \( \leq 10K \) DP entries?
    - yes: solve *optimally* using DPhyp
    - no: corner case?
      - yes: cannot linearize using GOO/DPhyp
      - no: \( \leq 100 \) relations?
        - yes: novel technique: *search space linearization* using linearized DP
        - no: gracefully introduce *greediness* to keep optimization time reasonable using GOO/linDP

Adaptive Optimization – How to Measure Complexity

<table>
<thead>
<tr>
<th>Structure</th>
<th>DP complexity</th>
<th>DP table size</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain</td>
<td>$O(n^3)$</td>
<td>$n^2$</td>
</tr>
<tr>
<td>clique</td>
<td>$O(3^n)$</td>
<td>$2^n$</td>
</tr>
</tbody>
</table>

- Complexity depends on the *structure* of the query graph
- *Size of DP table* as measure of complexity
- *Analyze query graph* to determine the size of the DP table
Adaptive Optimization – Small Queries

- Up to 10,000 DP entries
  - chains: up to 100 relations
  - cliques: less than 14 relations
- Run $DPhyp$
  - Adapts to the query graph’s structure
  - Completely and minimally enumerates all possibly optimal join orders without cross products
- Plan guaranteed to be optimal
- Optimization will be fast
Adaptive Optimization – Medium Queries

- Complexity depends on the *structure* of the query graph
- Can easily optimize *chain queries* on 100 relations exactly (polynomial runtime)
- Usually queries are not exactly linear
- Still benefit from this *fast optimization* through *search space linearization*
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known

R1 -- R2

R3 -- R4

R5 -- R6

R2 R1 R3 R5 R6 R4
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this linearization
- Optimally combine optimal solutions for subchains of increasing size

\[ R_1 \rightarrow R_2 \rightarrow R_3 \rightarrow R_4 \]

\[ R_1 \rightarrow R_2 \rightarrow R_5 \rightarrow R_6 \]

\[ R_2 \rightarrow R_1 \rightarrow R_3 \rightarrow R_5 \rightarrow R_6 \rightarrow R_4 \]
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this *linearization*
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this linearization
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this linearization
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this linearization
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this *linearization*
- Optimally combine optimal solutions for subchains of increasing size

Diagram:

```
R2  R1  R3  R5  R6  R4
```

```
R1 —— R2 —— R3
      \  |  /
       \ | /            \
        X  \   /
         \ /    \
          R5 —— R6
```
Join Ordering

Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this linearization
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this *linearization*
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this *linearization*
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this *linearization*
- Optimally combine optimal solutions for subchains of increasing size
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this *linearization*
- Optimally combine optimal solutions for subchains of increasing size
- But: how to know the optimal order?
Adaptive Optimization – Search Space Linearization

- Assume the order of relations in the optimal plan is known
- Polynomial DP algorithm to generate optimal plan from this *linearization*
- Optimally combine optimal solutions for subchains of increasing size
- But: how to know the optimal order?
- *IKKBZ* (TODS 3/’84, VLDB ’86): Optimal left-deep plan in $O(n^2)$
- Good alternative to the optimal relative order of relations
Adaptive Optimization – Linearized DP

Procedure

1. *Linearize* using IKKBZ
2. Build *best bushy plan* for linearization

Properties

- Runs in $O(n^3)$
- Result *at least as good as* the optimal left-deep plan
- With proper linearization, discovers globally *optimal bushy plan*
Adaptive Optimization – Large Queries

- Even linearized DP too expensive for the most complex queries
- Iterative Dynamic Programming (Kossmann & Stocker, TODS 1/2000):
  1. **Greedily build** query plan, e.g. using Greedy Operator Ordering (GOO)
  2. **Iteratively refine** by optimizing the most expensive sub trees of size $k$
     using DP
- Linearization greatly increases reordering freedom
  - originally: $k \approx 7$
  - linearized: $k = 100$
Generated Queries – Optimization Time

![Graph showing median optimization time over relations for different strategies: DPSize, DPSizeLinear, DPhyp, and adaptive. The x-axis represents the number of relations, ranging from 10 to 100. The y-axis represents the median optimization time in seconds, ranging from 0 to 40. The graph demonstrates that the adaptive strategy has the lowest median optimization time, followed by DPSizeLinear, and then DPhyp. The timeout is indicated by a skull and crossbones icon.]
Generated Queries – Optimization Time

- **timeout**: The graph indicates the time at which optimization fails to provide results.
- **DP**: The Disjunctive Programming curve shows a steady increase in optimization time with the number of relations.
- **linearizedDP**: The linearized DP curve is less steep compared to DP, indicating more efficient optimization.
- **minsel**: The minimum selection curve is the most efficient, showing minimal optimization time throughout.
- **genetic**: The genetic algorithm curve is initially efficient but increases significantly at higher numbers of relations.
- **IKKBZ**: The IKKBZ algorithm curve is similar to genetic, increasing significantly at higher numbers.
- **adaptive**: The adaptive optimization curve is efficient and maintains a low optimization time across all relation counts.

The graph visualizes the median optimization time [s] against the number of relations.
Generated Queries – Optimization Time

![Graph showing optimization time for different methods](image)

- **DP**
- **linearizedDP**
- **IKKBZ**
- **adaptive**
Generated Queries – Optimization Time

The graph illustrates the optimization time for generated queries with respect to the number of relations. The x-axis represents the number of relations, ranging from 10 to 1,000, while the y-axis represents the median optimization time in seconds. The adaptive method is shown in blue, minsel method in red, and genetic method in purple. The timeout is indicated by a gray line.
Generated Queries – Optimization Time

The diagram illustrates the optimization time for different methods as the number of relations increases. The x-axis represents the number of relations, ranging from 10 to 1,000, while the y-axis shows the median optimization time in seconds. The methods compared include DP, linearizedDP, minsel, IKKBZ, and genetic algorithms. The adaptive method is highlighted, showing its efficiency in optimizing generated queries over time.
Generated Queries – Plan Quality

Plan cost compared to cost of best plan found by any of the algorithms

Optimal plan known (371 queries)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>median</th>
<th>95%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPhyp</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Linearized DP</td>
<td>1.00</td>
<td>1.23</td>
<td>2.23</td>
</tr>
<tr>
<td>adaptive</td>
<td>1.00</td>
<td>1.10</td>
<td>2.23</td>
</tr>
</tbody>
</table>

- Most of the plans generated by linearized DP are optimal or near-optimal
- Adaptive Optimization additionally benefits from full DPhyp as long as it is fast
## Generated Queries – Plan Quality

### Linearized DP (≤ 100 relations; 1,000 queries)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>median</th>
<th>95%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IKKBZ</td>
<td>1.00</td>
<td>1.97</td>
<td>58.47</td>
</tr>
<tr>
<td>Linearized DP</td>
<td>1.00</td>
<td>1.12</td>
<td>2.57</td>
</tr>
<tr>
<td>adaptive</td>
<td>1.00</td>
<td>1.07</td>
<td>2.57</td>
</tr>
</tbody>
</table>

- DP phase in linearized DP significantly increases plan quality
## Generated Queries – Plan Quality

### Iterative Dynamic Programming (≤ 5,000 relations; 2,300 queries)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>median</th>
<th>95%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOO</td>
<td>1.05</td>
<td>2.81</td>
<td>19.18</td>
</tr>
<tr>
<td>GOO/DPhyp</td>
<td>1.01</td>
<td>2.53</td>
<td>19.18</td>
</tr>
<tr>
<td>GOO/linDP</td>
<td>1.00</td>
<td>1.60</td>
<td>4.02</td>
</tr>
<tr>
<td>adaptive</td>
<td>1.00</td>
<td>1.59</td>
<td>4.02</td>
</tr>
</tbody>
</table>

- Iterative DP benefits from additional freedom induced by linearized DP
- Adaptive Optimization generates good plans across the whole spectrum of queries
The algorithms so far have some drawbacks:

- greedy heuristics only heuristics
- will probably not find the optimal solution
- DP algorithms optimal, but very heavy weight
- especially memory consumption is high
- find a solution only after the complete search

Sometimes we want a more light-weight algorithm:

- low memory consumption
- stop if time runs out
- still find the optimal solution if possible
Generating Permutations (2)

We can achieve this when only considering left-deep trees:

- left-deep trees are permutations of the relations to be joined
- permutations can be generated directly
- generating all permutations is too expensive
- but some permutations can be ignored:
  Consider the join sequence $R_1 R_2 R_3 R_4$. If we know that $R_1 R_3 R_2$ is cheaper than $R_1 R_2 R_3$, we do not have to consider $R_1 R_2 R_3 R_4$.

Idea: successively add a relation. An extended sequence is only explored if exchanging the last two relations does not result in a cheaper sequence.
Recursive Search

ConstructPermutations($R$)

**Input:** a set of relations $R = \{R_1, \ldots, R_n\}$ to be joined

**Output:** an optimal left-deep join tree

$B = \epsilon$

$P = \epsilon$

for each $R_i \in R$ {
    ConstructPermutationsRec($P \circ < R_i >, R \setminus \{R_i\}, B$)
}

return $B$

- algorithm considers a prefix $P$ and the rest $R$
- keeps track of the best tree found so far $B$
- increases the prefix recursively
Recursive Search (2)

ConstructPermutationsRec($P$, $R$, $B$)

**Input:** a prefix $P$, remaining relations $R$, best plan $B$

**Output:** side effects on $B$

if $|R| = 0$ {
    if $B = \epsilon \lor C(B) > C(P)$ {
        $B = P$
    }
} else {
    for each $R_i \in R$ {
        if $C(P \circ < R_i >) \leq C(P[1 : |P| - 1] \circ < R_i, P[|P|] >)$ {
            ConstructPermutationsRec($P \circ < R_i >$, $R \setminus \{R_i\}$, $B$)
        }
    }
}
Remarks

Good:

- linear memory
- immediately produces plan alternatives
- anytime algorithm
- finds the optimal plan eventually

Bad:

- worst-case runtime if ties occur
- worst-case runtime if no ties occur is an open problem

Often fast, can be stopped anytime, but may perform poorly.
Transformative Approaches

Main idea: [7]

- use equivalences directly (associativity, commutativity)
- would make integrating new equivalences easy

Problems:

- how to navigate the search space
- equivalences have no order
- how to guarantee finding the optimal solution
- how to avoid exhaustive search
Rule Set

\[ R_1 \bowtie R_2 \quad \leadsto \quad R_2 \bowtie R_1 \quad \text{Commutativity} \]
\[ (R_1 \bowtie R_2) \bowtie R_3 \quad \leadsto \quad R_1 \bowtie (R_2 \bowtie R_3) \quad \text{Right Associativity} \]
\[ R_1 \bowtie (R_2 \bowtie R_3) \quad \leadsto \quad (R_1 \bowtie R_2) \bowtie R_3 \quad \text{Left Associativity} \]
\[ (R_1 \bowtie R_2) \bowtie R_3 \quad \leadsto \quad (R_1 \bowtie R_3) \bowtie R_2 \quad \text{Left Join Exchange} \]
\[ R_1 \bowtie (R_2 \bowtie R_3) \quad \leadsto \quad R_2 \bowtie (R_1 \bowtie R_3) \quad \text{Right Join Exchange} \]

Two more rules are often used to transform left-deep trees:

- \textit{swap} exchanges two arbitrary relations in a left-deep tree
- \textit{3Cycle} performs a cyclic rotation of three arbitrary relations in a left-deep tree.

To try another join method, another rule called \textit{join method exchange} is introduced.
Rule Set RS-0

- commutativity
- left-associativity
- right-associativity
Basic Algorithm

ExhaustiveTransformation($\{R_1, \ldots, R_n\}$)

**Input:** a set of relations  
**Output:** an optimal join tree

Let $T$ be an arbitrary join tree for all relations  
Done = $\emptyset$ // contains all trees processed  
ToDo = $\{T\}$ // contains all trees to be processed

**while** $|\text{ToDo}| > 0$ {

    $T =$ an arbitrary tree in ToDo  
    ToDo = ToDo \ $T$;  
    Done = Done $\cup$ $\{T\}$;  
    Trees = ApplyTransformations($T$);  
    **for each** $T \in$ Trees {
        **if** $T \notin$ ToDo $\cup$ Done  
            ToDo = ToDo $\cup$ $\{T\}$
    }
}

**return** $\text{arg min}_{T \in \text{Done}} C(T)$
Basic Algorithm (2)

ApplyTransformations($T$)

**Input:** join tree

**Output:** all trees derivable by associativity and commutativity

$Trees = \emptyset$

$Subtrees =$ all subtrees of $T$ rooted at inner nodes

**for each** $S \in Subtrees$ {

if $S$ is of the form $S_1 \bowtie S_2$

$Trees = Trees \cup \{S_2 \bowtie S_1\}$

if $S$ is of the form $(S_1 \bowtie S_2) \bowtie S_3$

$Trees = Trees \cup \{S_1 \bowtie (S_2 \bowtie S_3)\}$

if $S$ is of the form $S_1 \bowtie (S_2 \bowtie S_3)$

$Trees = Trees \cup \{(S_1 \bowtie S_2) \bowtie S_3\}$

}

**return** $Trees$;
Remarks

- if no cross products are to be considered, extend if conditions for associativity rules.
- problem 1: explores the whole search space
- problem 2: generates join trees more than once
- problem 3: sharing of subtrees is non-trivial
Search Space
Introducing the Memo Structure

A memoization strategy is used to keep the runtime reasonable:

• for any subset of relations, dynamic programming remembers the best join tree.
• this does not quite suffice for the transformation-based approach.
• instead, we have to keep all join trees generated so far including those differing in the order of the arguments of a join operator.
• however, subtrees can be shared.
• this is done by keeping pointers into the data structure (see next slide).
## Memo Structure Example

<table>
<thead>
<tr>
<th>{R_1, R_2, R_3}</th>
<th>{R_1, R_2} \times R_3, R_3 \times {R_1, R_2}, {R_1, R_3} \times R_2, R_2 \times {R_1, R_3}, {R_2, R_3} \times R_1, R_1 \times {R_2, R_3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>{R_2, R_3}</td>
<td>{R_2} \times {R_3}, {R_3} \times {R_2}</td>
</tr>
<tr>
<td>{R_1, R_3}</td>
<td>{R_1} \times {R_3}, {R_3} \times {R_1}</td>
</tr>
<tr>
<td>{R_1, R_2}</td>
<td>{R_1} \times {R_2}, {R_2} \times {R_1}</td>
</tr>
<tr>
<td>{R_3}</td>
<td>R_3</td>
</tr>
<tr>
<td>{R_2}</td>
<td>R_2</td>
</tr>
<tr>
<td>{R_1}</td>
<td>R_1</td>
</tr>
</tbody>
</table>

- in Memo Structure: arguments are pointers to classes
- Algorithm: `ExploreClass` expands a class
- Algorithm: `ApplyTransformation2` expands a member of a class
Memoizing Algorithm

ExhaustiveTransformation2(Query Graph $G$)

**Input:** a query specification for relations $\{R_1, \ldots, R_n\}$.

**Output:** an optimal join tree

initialize MEMO structure

ExploreClass($\{R_1, \ldots, R_n\}$)

return $\arg\min_{T \in \text{class } \{R_1, \ldots, R_n\}} C(T)$

- stored an arbitrary join tree in the memo structure
- explores alternatives recursively
Memoizing Algorithm (2)

ExploreClass($C$)

**Input:** a class $C \subseteq \{R_1, \ldots, R_n\}$

**Output:** none, but has side-effect on MEMO-structure

while not all join trees in $C$ have been explored {
    choose an unexplored join tree $T$ in $C$
    ApplyTransformation2($T$)
    mark $T$ as explored
}

- considers all alternatives within one class
- transformations themselves are done in ApplyTransformation2
Memoizing Algorithm (3)

ApplyTransformations2( $T$ )

**Input:** a join tree of a class $C$

**Output:** none, but has side-effect on MEMO-structure

ExploreClass(left-child($T$))
ExploreClass(right-child($T$));

for each transformation $T$ and class member of child classes {
    for each $T'$ resulting from applying $T$ to $T$ {
        if $T'$ not in MEMO structure {
            add $T'$ to class $C$ of MEMO structure
        }
    }
}

- first explores subtrees
- then applies all known transformations to the tree
- stores new trees in the memo structure
Remarks

• Applying ExhaustiveTransformation2 with a rule set consisting of Commutativity and Left and Right Associativity generates $4^n - 3^{n+1} + 2^{n+2} - n - 2$ duplicates

• Contrast this with the number of join trees contained in a completely filled MEMO structure: $3^n - 2^{n+1} + n + 1$

• Solve the problem of duplicate generation by disabling applied rules.
Rule Set RS-1

$T_1$: Commutativity \[ C_1 \Join_0 C_2 \leadsto C_2 \Join_1 C_1 \]
Disable all transformations $T_1$, $T_2$, and $T_3$ for $\Join_1$.

$T_2$: Right Associativity \[ (C_1 \Join_0 C_2) \Join_1 C_3 \leadsto C_1 \Join_2 (C_2 \Join_3 C_3) \]
Disable transformations $T_2$ and $T_3$ for $\Join_2$ and enable all rules for $\Join_3$.

$T_3$: Left associativity \[ C_1 \Join_0 (C_2 \Join_1 C_3) \leadsto (C_1 \Join_2 C_2) \Join_3 C_3 \]
Disable transformations $T_2$ and $T_3$ for $\Join_3$ and enable all rules for $\Join_2$. 
Example for chain $R_1 - R_2 - R_3 - R_4$

<table>
<thead>
<tr>
<th>Class</th>
<th>Initialization</th>
<th>Transformation</th>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>${R_1, R_2, R_3, R_4}$</td>
<td>${R_1, R_2} \bowtie_{111} {R_3, R_4}$</td>
<td>${R_3, R_4} \bowtie_{000} {R_1, R_2}$</td>
<td>3</td>
</tr>
<tr>
<td>${R_2, R_3, R_4}$</td>
<td></td>
<td>$R_1 \bowtie_{100} {R_2, R_3, R_4}$</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>${R_1, R_2, R_3} \bowtie_{100} R_4$</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>${R_2, R_3, R_4} \bowtie_{000} R_1$</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R_4 \bowtie_{000} {R_1, R_2, R_3}$</td>
<td>10</td>
</tr>
<tr>
<td>${R_1, R_3, R_4}$</td>
<td></td>
<td>$R_2 \bowtie_{111} {R_3, R_4}$</td>
<td>4</td>
</tr>
<tr>
<td>${R_1, R_2, R_4}$</td>
<td></td>
<td>${R_3, R_4} \bowtie_{000} R_2$</td>
<td>6</td>
</tr>
<tr>
<td>${R_1, R_2, R_3}$</td>
<td></td>
<td>${R_2, R_3} \bowtie_{100} R_4$</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R_4 \bowtie_{000} {R_2, R_3}$</td>
<td>7</td>
</tr>
<tr>
<td>${R_3, R_4}$</td>
<td></td>
<td>${R_1, R_2} \bowtie_{111} R_3$</td>
<td>5</td>
</tr>
<tr>
<td>${R_2, R_4}$</td>
<td></td>
<td>$R_3 \bowtie_{000} {R_1, R_2}$</td>
<td>9</td>
</tr>
<tr>
<td>${R_2, R_3}$</td>
<td></td>
<td>$R_1 \bowtie_{100} {R_2, R_3}$</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>${R_2, R_3} \bowtie_{000} R_1$</td>
<td>9</td>
</tr>
<tr>
<td>${R_1, R_4}$</td>
<td></td>
<td>$R_3 \bowtie_{111} R_4$</td>
<td>2</td>
</tr>
<tr>
<td>${R_1, R_3}$</td>
<td></td>
<td>$R_4 \bowtie_{000} R_3$</td>
<td></td>
</tr>
<tr>
<td>${R_1, R_2}$</td>
<td></td>
<td>$R_1 \bowtie_{111} R_2$</td>
<td>1</td>
</tr>
<tr>
<td>${R_1, R_3}$</td>
<td></td>
<td>$R_2 \bowtie_{000} R_1$</td>
<td>1</td>
</tr>
</tbody>
</table>
Rule Set RS-2

Bushy Trees: Rule set for clique queries and if cross products are allowed:

$T_1$: Commutativity  $C_1 \bowtie_0 C_2 \leadsto C_2 \bowtie_1 C_1$
_Disable all transformations $T_1$, $T_2$, $T_3$, and $T_4$ for $\bowtie_1$.

$T_2$: Right Associativity  $(C_1 \bowtie_0 C_2) \bowtie_1 C_3 \leadsto C_1 \bowtie_2 (C_2 \bowtie_3 C_3)$
_Disable transformations $T_2$, $T_3$, and $T_4$ for $\bowtie_2$.

$T_3$: Left Associativity  $C_1 \bowtie_0 (C_2 \bowtie_1 C_3) \leadsto (C_1 \bowtie_2 C_2) \bowtie_3 C_3$
_Disable transformations $T_2$, $T_3$ and $T_4$ for $\bowtie_3$.

$T_4$: Exchange  $(C_1 \bowtie_0 C_2) \bowtie_1 (C_3 \bowtie_2 C_4) \leadsto (C_1 \bowtie_3 C_3) \bowtie_4 (C_2 \bowtie_5 C_4)$
_Disable all transformations $T_1$, $T_2$, $T_3$, and $T_4$ for $\bowtie_4$.

If we initialize the MEMO structure with left-deep trees, we can strip down the above rule set to Commutativity and Left Associativity. Reason: from a left-deep join tree we can generate all bushy trees with only these two rules.
Rule Set RS-3

Left-deep trees:

\( T_1 \) Commutativity \( \: R_1 \bowtie_{0} R_2 \sim \rightarrow R_2 \bowtie_{1} R_1 \)

Here, the \( R_i \) are restricted to classes with exactly one relation. \( T_1 \) is disabled for \( \bowtie_{1} \).

\( T_2 \) Right Join Exchange \( \: ( C_1 \bowtie_{0} C_2 ) \bowtie_{1} C_3 \sim \rightarrow ( C_1 \bowtie_{2} C_3 ) \bowtie_{3} C_2 \)

Disable \( T_2 \) for \( \bowtie_{3} \).
Generating Random Join Trees

Generating a random join tree is quite useful:

- allows for cost sampling
- randomized optimization procedures
- basis for Simulated Annealing, Iterative Improvement etc.
- easy with cross products, difficult without
- we consider with cross products first

Main problems:

- generating all join trees (potentially)
- creating all with the same probability
Ranking/Unranking

Let $S$ be a set with $n$ elements.

- a bijective mapping $f : S \rightarrow [0, n]$ is called ranking
- a bijective mapping $f : [0, n] \rightarrow S$ is called unranking

Given an unranking function, we can generate random elements in $S$ by generating a random number in $[0, n]$ and unranking this number.

Challenge: making unranking fast.
Random Permutations

Every permutation corresponds to a left-deep join tree possibly with cross products.
Standard algorithm to generate random permutations is the starting point for the algorithm:

\[
\text{for each } k \in [0, n[ \text{ descending} \\
\quad \text{swap}(\pi[k], \pi[\text{random}(k)])
\]

Array \( \pi \) initialized with elements \([0, n[. \text{ random}(k) \) generates a random number in \([0, k]). \)
Random Permutations

- Assume the random elements produced by the algorithm are $r_{n-1}, \ldots, r_0$ where $0 \leq r_i \leq i$.

- Thus, there are exactly $n(n-1)(n-2)\ldots1 = n!$ such sequences and there is a one to one correspondence between these sequences and the set of all permutations.

- Unrank $r \in [0, n!]$ by turning it into a unique sequence of values $r_{n-1}, \ldots, r_0$.
  Note that after executing the swap with $r_{n-1}$ every value in $[0, n]$ is possible at position $\pi[n-1]$.
  Further, $\pi[n-1]$ is never touched again.

- Hence, we can unrank $r$ as follows. We first set $r_{n-1} = r \mod n$ and perform the swap. Then, we define $r' = \lfloor r/n \rfloor$ and iteratively unrank $r'$ to construct a permutation of $n-1$ elements.
Generating Random Permutations

Unrank($n, r$)

**Input:** the number $n$ of elements to be permuted and the rank $r$ of the permutation to be constructed

**Output:** a permutation $\pi$

for each $0 \leq i < n$

$\pi[i] = i$

for each $n \geq i > 0$ descending {

swap($\pi[i - 1], \pi[r \mod i]$)

$r = \lfloor r/i \rfloor$

}

return $\pi$;
Generating Random Bushy Trees with Cross Products

Steps of the algorithm:

1. Generate a random number $b$ in $[0, C(n)]$.
2. Unrank $b$ to obtain a bushy tree with $n – 1$ inner nodes.
3. Generate a random number $p$ in $[0, n!]$.
4. Unrank $p$ to obtain a permutation.
5. Attach the relations in order $p$ from left to right as leaf nodes to the binary tree obtained in Step 2.

The only step that we have still to discuss is Step 2.
Tree Encoding

- **Preorder traversal:**
  - Inner node: ’(’
  - Leaf Node: ’)’
  
  Skip last leaf node.
- Replace ’(’ by 1 and ’)’ by 0
- Just take positions of 1s.

Example: all trees with four inner nodes:
- The ranks are in \([0, 14[\)
Tree Ranking Example
Unranking Binary Trees

We establish a bijection between Dyck words and paths in a grid:

Every path from \((0, 0)\) to \((2n, 0)\) uniquely corresponds to a Dyck word.
Counting Paths

The number of different paths from $(0, 0)$ to $(i, j)$ can be computed by

$$p(i, j) = \frac{j + 1}{i + 1} \left( \frac{i + 1}{\frac{1}{2}(i + j) + 1} \right)$$

These numbers are the *Ballot numbers*. The number of paths from $(i, j)$ to $(2n, 0)$ can thus be computed as:

$$q(i, j) = p(2n - i, j)$$

Note the special case $q(0, 0) = p(2n, 0) = C(n)$. 
Unranking Outline

- We open a parenthesis (go from \((i, j)\) to \((i + 1, j + 1)\)) as long as the number of paths from that point does no longer exceed our rank \(r\).
- If it does, we close a parenthesis (go from \((i, j)\) to \((i + 1, j - 1)\)).
- Assume, that we went upwards to \((i, j)\) and then had to go down to \((i + 1, j - 1)\).
  We subtract the number of paths from \((i + 1, j + 1)\) from our rank \(r\) and proceed iteratively from \((i + 1, j - 1)\) by going up as long as possible and going down again.
- Remembering the number of parenthesis opened and closed along our way results in the required encoding.
Generating Bushy Trees

UnrankTree\((n, r)\)

**Input:** a number of inner nodes \(n\) and a rank \(r \in [0, C(n)]\)

**Output:** encoding of the inner leafes of a tree

open = 1, close = 0

pos = 1, encoding = \(<1>\)

**while** \(|\text{encoding}| < n\) {

\(k = q(\text{open}+\text{close}, \text{open}-\text{close})\)

**if** \(k \leq r\) {

\(r = r - k, \text{close} = \text{close} + 1\)

}**else** {

encoding = encoding \(\circ \,<\text{pos}\,>, \text{open} = \text{open} + 1\)

}

pos = pos + 1

}**return** encoding
Generating Random Trees Without Cross Products

Tree queries only!

- query graph: $G = (V, E)$, $|V| = n$, $G$ must be a tree.
- level: root has level 0, children thereof 1, etc.
- $T_G$: join trees for $G$

[8]
Partitioning $\mathcal{T}_G$

$\mathcal{T}_G^{v(k)} \subseteq \mathcal{T}_G$: subset of join trees where the leaf node (i.e. relation) $v$ occurs at level $k$.

Observations:

- $n = 1$: $|\mathcal{T}_G| = |\mathcal{T}_G^{v(0)}| = 1$
- $n > 1$: $|\mathcal{T}_G^{v(0)}| = 0$ (top is a join and no relation)
- The maximum level that can occur in any join tree is $n - 1$.
  Hence: $|\mathcal{T}_G^{v(k)}| = 0$ if $k \geq n$.
- $\mathcal{T}_G = \bigcup_{k=0}^{n} \mathcal{T}_G^{v(k)}$
- $\mathcal{T}_G^{v(i)} \cap \mathcal{T}_G^{v(j)} = \emptyset$ for $i \neq j$
- Thus: $|\mathcal{T}_G| = \sum_{k=0}^{n} |\mathcal{T}_G^{v(k)}|$
The Specification

- The algorithm will generate an unordered tree with $n$ leaf nodes.
- If we wish to have a random ordered tree, we have to pick one of the $2^{n-1}$ possibilities to order the $(n - 1)$ joins within the tree.
The Procedure

1. List merges (notation, specification, counting, unranking)
2. Join tree construction: leaf-insertion and tree-merging
3. Standard Decomposition Graph (SDG): describes all valid join trees
4. Counting
5. Unranking algorithm
List Merge

- Lists: Prolog-Notation: \(< a | t >\)
- Property \(P\) on elements
- A list \(l'\) is the *projection* of a list \(L\) on \(P\), if \(L'\) contains all elements of \(L\) satisfying the property \(P\).
  Thereby, the order is retained.
- A list \(L\) is a *merge* of two disjoint lists \(L_1\) and \(L_2\), if \(L\) contains all elements from \(L_1\) and \(L_2\) and both are projections of \(L\).
Example

(R, S, [1, 1, 0])  (R, S, [2, 0, 0])  (R, S, [0, 2, 0])
List Merge: Specification

A merge of a list $L_1$ with a list $L_2$ whose respective lengths are $l_1$ and $l_2$ can be described by an array $\alpha = [\alpha_0, \ldots, \alpha_{l_2}]$ of non-negative integers whose sum is equal to $l_1$, i.e. $\sum_{i=0}^{l_2} \alpha_i = |l_1|$.

- We obtain the merged list $L$ by first taking $\alpha_0$ elements from $L_1$.
- Then, an element from $L_2$ follows. Then follow $\alpha_1$ elements from $L_1$ and the next element of $L_2$ and so on.
- Finally follow the last $\alpha_{l_2}$ elements of $L_1$. 
List Merge: Counting

Non-negative integer decomposition:

- What is the number of decompositions of a non-negative integer $n$ into $k$ non-negative integers $\alpha_i$ with $\sum_{i=1}^{k} \alpha_k = n$.

Answer: $\binom{n+k-1}{k-1}$
List Merge: Counting (2)

Since we have to decompose $l_1$ into $l_2 + 1$ non-negative integers, the number of possible merges is $M(l_1, l_2) = \binom{l_1 + l_2}{l_2}$.

The observation $M(l_1, l_2) = M(l_1 - 1, l_2) + M(l_1, l_2 - 1)$ allows us to construct an array of size $n \times n$ in $O(n^2)$ that materializes the values for $M$. This array will allow us to rank list merges in $O(l_1 + l_2)$. 
List Merge: Unranking: General Idea

The idea for establishing a bijection between $[1, M(l_1, l_2)]$ and the possible $\alpha$s is a general one and used for all subsequent algorithms of this section. Assume we want to rank the elements of some set $S$ and $S = \bigcup_{i=0}^{n} S_i$ is partitioned into disjoint $S_i$.

1. If we want to rank $x \in S_k$, we first find the local rank of $x \in S_k$.
2. The rank of $x$ is then $\sum_{i=0}^{k-1} |S_i| + \text{local-rank}(x, S_k)$.
3. To unrank some number $r \in [1, N]$, we first find $k$ such that $k = \min_j r \leq \sum_{i=0}^{j} |S_i|$.
4. We proceed by unranking with the new local rank $r' = r - \sum_{i=0}^{k-1} |S_i|$ within $S_k$. 
List Merge: Unranking

We partition the set of all possible merges into subsets.

- Each subset is determined by $\alpha_0$. For example, the set of possible merges of two lists $L_1$ and $L_2$ with length $l_1 = l_2 = 4$ is partitioned into subsets with $\alpha_0 = j$ for $0 \leq j \leq 4$.
- In each partition, we have $M(l_1 - j, l_2 - 1)$ elements.
- To unrank a number $r \in [1, M(l_1, l_2)]$ we first determine the partition by computing $k = \min_j r \leq \sum_{i=0}^j M(j, l_2 - 1)$. Then, $\alpha_0 = l_1 - k$.
- With the new rank $r' = r - \sum_{i=0}^k M(j, l_2 - 1)$, we start iterating all over.
Example

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\alpha_0$</th>
<th>$(k, l_2 - 1)$</th>
<th>$M(k, l_2 - 1)$</th>
<th>rank intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>(0, 3)</td>
<td>1</td>
<td>[1, 1]</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>(1, 3)</td>
<td>4</td>
<td>[2, 5]</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>(2, 3)</td>
<td>10</td>
<td>[6, 15]</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>(3, 3)</td>
<td>20</td>
<td>[16, 35]</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>(4, 3)</td>
<td>35</td>
<td>[36, 70]</td>
</tr>
</tbody>
</table>
UnrankDecomposition\((r, l_1, l_2)\)

**Input:** a rank \(r\), two list sizes \(l_1\) and \(l_2\)

**Output:** encoding of the inner leafes of a tree

\[
\alpha = <>, \quad k = 0
\]

**while** \(l_1 > 0 \land l_2 > 0\) {
  \[
  m = M(k, l_2 - 1)
  \]
  **if** \(r \leq m\) {
    \[
    \alpha = \alpha e < l_1 - k >
    \]
    \[
    l_1 = k, \quad k = 0, \quad l_2 = l_2 - 1
    \]
  } **else** {
    \[
    r = r - m, \quad k = k + 1
    \]
  }
}

**return** \(\alpha < l_1 > \circ \bigcirc |\alpha|_{1 \leq i < l_2} < 0 >\)
Anchored List Representation of Join Trees

**Definition** Let $T$ be a join tree and $v$ be a leaf of $T$. The *anchored list representation* $L$ of $T$ is constructed as follows:

- If $T$ consists of the single leaf node $v$, then $L = \langle \rangle$.
- If $T = (T_1 \Join T_2)$ and without loss of generality $v$ occurs in $T_2$, then $L = \langle T_1 | L_2 \rangle$ where $L_2$ is the anchored list representation of $T_2$.

We then write $T = (L, v)$.

**Observation** If $T = (L, v) \in \mathcal{T}_G$ then $T \in \mathcal{T}_G^v(k) \iff |L| = k$
Leaf-Insertion: Example
Leaf-Insertion

**Definition** Let $G = (V, E)$ be a query graph, $T$ a join tree of $G$. $v \in V$ be such that $G' = G|_{V \setminus \{v\}}$ is connected, $(v, w) \in E$, $1 \leq k < n$, and

$$T = (\langle T_1, \ldots, T_{k-1}, v, T_{k+1}, \ldots, T_n \rangle, w)$$
$$T' = (\langle T_1, \ldots, T_{k-1}, T_{k+1}, \ldots, T_n \rangle, w).$$

Then we call $(T', k)$ an *insertion pair* on $v$ and say that $T$ is *decomposed into* (or *constructed from*) the pair $(T', k)$ on $v$.

**Observation:** Leaf-insertion defines a bijective mapping between $T_T^{v(k)}_G$ and insertion pairs $(T', k)$ on $v$, where $T'$ is an element of the disjoint union $\bigcup_{i=k-1}^{n-2} T_{G'}^w(i)$. 
Tree-Merging: Example

(R, S, [1, 1, 0])  (R, S, [2, 0, 0])  (R, S, [0, 2, 0])
Tree-Merging

Two trees $R = (L_R, w)$ and $S = (L_S, w)$ on a common leaf $w$ are merged by merging their anchored list representations.

**Definition.** Let $G = (V, E)$ be a query graph, $w \in V$, $T = (L, w)$ a join tree of $G$, $V_1, V_2 \subseteq V$ such that $G_1 = G|_{V_1}$ and $G_2 = G|_{V_2}$ are connected, $V_1 \cup V_2 = V$, and $V_1 \cap V_2 = \{w\}$. For $i = 1, 2$:

- Define the property $P_i$ to be “every leaf of the subtree is in $V_i$”,
- Let $L_i$ be the projection of $L$ on $P_i$.
- $T_i = (L_i, w)$.

Let $\alpha$ be the integer decomposition such that $L$ is the result of merging $L_1$ and $L_2$ on $\alpha$. Then, we call $(T_1, T_2, \alpha)$ a merge triplet. We say that $T$ is *decomposed into* (constructed from) $(T_1, T_2, \alpha)$ on $V_1$ and $V_2$. 
Observation

Tree-Merging defines a bijective mapping between $T_G^{w(k)}$ and merge triplets $(T_1, T_2, \alpha)$, where $T_1 \in T_{G_1}^{w(i)}$, $T_2 \in T_{G_2}^{w(k-i)}$, and $\alpha$ specifies a merge of two lists of sizes $i$ and $k-i$. Further, the number of these merges (i.e. the number of possibilities for $\alpha$) is $\binom{i+(k-i)}{k-i} = \binom{k}{i}$. 
A *standard decomposition graph* of a query graph describes the possible constructions of join trees. It is not unique (for \( n > 1 \)) but anyone can be used to construct all possible unordered join trees. For each of our two operations it has one kind of inner nodes:

- A unary node labeled \( +_v \) stands for leaf-insertion of \( v \).
- A binary node labeled \( \ast_w \) stands for tree-merging its subtrees whose only common leaf is \( w \).
Constructing a Standard Decomposition Graph

The standard decomposition graph of a query graph $G = (V, E)$ is constructed in three steps:

1. pick an arbitrary node $r \in V$ as its root node
2. transform $G$ into a tree $G'$ by directing all edges away from $r$;
3. call $QG2SDG(G', r)$
Constructing a Standard Decomposition Graph (2)

QG2SDG(G', r)

**Input:** a query tree \( G' = (V, E) \) and its root \( r \)

**Output:** a standard query decomposition tree of \( G' \)

Let \( \{w_1, \ldots, w_n\} \) be the children of \( v \)

switch \( n \) {

  case 0: label \( v \) with "v"

  case 1:
    label \( v \) as "+v"
    QG2SDG(G', w_1)

  otherwise:
    label \( v \) as "*v"
    create new nodes \( l, r \) with label \( +v \)
    \( E = E \setminus \{(v, w_i) | 1 \leq i \leq n\} \)
    \( E = E \cup \{(v, l), (v, r), (l, w_1)\} \cup \{(r, w_i) | 2 \leq i \leq n\} \)
    QG2SDG(G', l), QG2SDG(G', r)

} return \( G' \)
Constructing a Standard Decomposition Graph (3)
Counting

For efficient access to the number of join trees in some partition $\mathcal{T}_G^{v(k)}$ in the unranking algorithm, we materialize these numbers. This is done in the count array.

The semantics of a count array $[c_0, c_1, \ldots, c_n]$ of a node $u$ with label $\circ_v$ ($\circ \in \{+, *\}$) of the SDG is that

- $u$ can construct $c_i$ different trees in which leaf $v$ is at level $i$.

Then, the total number of trees for a query can be computed by summing up all the $c_i$ in the count array of the root node of the decomposition tree.
Counting (2)

To compute the count and an additional summand adornment of a node labeled $+_v$, we use the following lemma:

**Lemma.** Let $G = (V, E)$ be a query graph with $n$ nodes, $v \in V$ such that $G' = G|_{V \setminus v}$ is connected, $(v, w) \in E$, and $1 \leq k < n$. Then

$$|T_G^v(k)| = \sum_{i \geq k-1} |T_{G'}^w(i)|$$
Counting (3)

The sets $\mathcal{T}_G^{v(i)}$ used in the summands of the former Lemma directly correspond to subsets $\mathcal{T}_G^{v(k),i}$ ($k - 1 \leq i \leq n - 2$) defined such that $T \in \mathcal{T}_G^{v(k),i}$ if

1. $T \in \mathcal{T}_G^{v(k)}$,
2. the insertion pair on $v$ of $T$ is $(T', k)$, and
3. $T' \in \mathcal{T}_G^{w(i)}$.

Further, $|\mathcal{T}_G^{v(k),i}| = |\mathcal{T}_G^{w(i)}|$. For efficiency, we materialize the summands in an array of arrays summands.
Counting (4)

To compute the count and summand adornment of a node labeled $*_v$, we use the following lemma.

**Lemma.** Let $G = (V, E)$ be a query graph, $w \in V$, $T = (L, w)$ a join tree of $G$, $V_1, V_2 \subseteq V$ such that $G_1 = G|_{V_1}$ and $G_2 = G|_{V_2}$ are connected, $V_1 \cup V_2 = V$, and $V_1 \cap V_2 = \{v\}$. Then

$$|T_G^{v(k)}| = \sum_{i} \binom{k}{i} |T_{G_1}^{v(i)}||T_{G_2}^{v(k-i)}|$$
Counting (5)

The sets $\mathcal{T}_{G}^{-w(i)}$ used in the summands of the previous Lemma directly correspond to subsets $\mathcal{T}_{G}^{v(k),i} (0 \leq i \leq k)$ defined such that $T \in \mathcal{T}_{G}^{v(k),i}$ if

1. $T \in \mathcal{T}_{G}^{v(k)}$,
2. the merge triplet on $V_1$ and $V_2$ of $T$ is $(T_1, T_2, \alpha)$, and
3. $T_1 \in \mathcal{T}_{G_1}^{v(i)}$.

Further, $|\mathcal{T}_{G}^{v(k),i}| = \binom{k}{i} |\mathcal{T}_{G_1}^{v(i)}| |\mathcal{T}_{G_2}^{v(k-i)}|$. 
Counting (6)

**Observation:** Assume a node \( v \) whose count array is \( [c_1, \ldots, c_m] \) and whose summands is \( s = [s^0, \ldots, s^n] \) with \( s_i = [s^i_0, \ldots, s^i_m] \), then

\[
c_i = \sum_{j=0}^{m} s^i_j
\]

holds.

The following algorithm has worst-case complexity \( O(n^3) \).

Looking at the count array of the root node of the following SDG, we see that the total number of join trees for our example query graph is 18.
SDG example

\( a \rightarrow b \rightarrow c \rightarrow d \)

\( e \)

\( a \leftarrow b \leftarrow c \leftarrow d \)

\( e \)

\( c \leftarrow b \leftarrow d \)

\( [0, 1, 1] \)

\( b \) +

\( +_e [0, 5, 5, 5, 3] \)

\( c \) *

\( *_c [0, 0, 2, 3] \)

\( +c [0, 1] \)

\( d \) +

\( \) +

\( a \)

\( [1] \)

\( [0, 1] \)

\( b \) +

\( +b [0, 1] \)

\( a \)
Annotating the SDG

Adorn($v$)

**Input:** a node $v$ of the SDG

**Output:** $v$ and nodes below are adorned by count and summands

Let $\{w_1, \ldots, w_n\}$ be the children of $v$

```
switch (n) {
  case 0: count($v$) = [1] // no summands for $v$
  case 1: Adorn($w_1$)
    assume count($w_1$) = [$c_1^0, \ldots, c_1^{m_1}$];
    count($v$) = [0, $c_1, \ldots, c_{m_1 + 1}$] where $c_k = \sum_{i=k-1}^{m_1} c_i^1$
    summands($v$) = [$s^0, \ldots, s^{m_1 + 1}$] where $s^k = [s_0^k, \ldots, s_{m_1 + 1}^k]$ and
    $s_i^k = \begin{cases} 
    c_i^1 & \text{if } 0 < k \text{ and } k - 1 \leq i \\
    0 & \text{else}
    \end{cases}$
```
Join Ordering  Randomized Approaches

Annotating the SDG (2)

case 2:
Adorn\(w_1\)
Adorn\(w_2\)
assume count\(w_1\) = \([c^1_0, \ldots, c^1_{m_1}]\)
assume count\(w_2\) = \([c^2_0, \ldots, c^2_{m_2}]\)
count\(v\) = \([c_0, \ldots, c_{m_1+m_2}]\) where
\[c_k = \sum_{i=0}^{m_1} \binom{k}{i} c^1_i c^2_{k-i}; \quad // \quad c^2_i = 0 \text{ for } i \not\in \{0, \ldots, m_2\}\]
summands\(v\) = \([s^0, \ldots, s^{m_1+m_2}]\) where \(s^k = [s^k_0, \ldots, s^k_{m_1}]\) and
\[s^k_i = \begin{cases} \binom{k}{i} c^1_i c^2_{k-i} & \text{if } 0 \leq k - i \leq m_2 \\ 0 & \text{else} \end{cases} \]
Unranking: top-level procedure

The algorithm `UnrankLocalTreeNoCross` called by `UnrankTreeNoCross` adorns the standard decomposition graph with insert-at and merge-using annotations. These can then be used to extract the join tree.

`UnrankTreeNoCross(r, v)`

**Input:** a rank $r$ and the root $v$ of the SDG

**Output:** adorned SDG

let $\text{count}(v) = [x_0, \ldots, x_m]$  

$k = \min_j r \leq \sum_{i=0}^{j} x_i$  

$r' = r - \sum_{i=0}^{k-1} x_i$  

`UnrankLocalTreeNoCross(v, r', k)`
Unranking: Example

The following table shows the intervals associated with the partitions $T^e_k$ for our standard decomposition graph:

<table>
<thead>
<tr>
<th>Partition</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^e_G(1)$</td>
<td>[1, 5]</td>
</tr>
<tr>
<td>$T^e_G(2)$</td>
<td>[6, 10]</td>
</tr>
<tr>
<td>$T^e_G(3)$</td>
<td>[11, 15]</td>
</tr>
<tr>
<td>$T^e_G(4)$</td>
<td>[16, 18]</td>
</tr>
</tbody>
</table>
Unranking: the last utility function

The unranking procedure makes use of unranking decompositions and unranking triples. For the latter and a given $X, Y, Z$, we need to assign each member in

$$\{(x, y, z) | 1 \leq x \leq X, 1 \leq y \leq Y, 1 \leq z \leq Z\}$$

a unique number in $[1, XYZ]$ and base an unranking algorithm on this assignment. We call the function $\text{UnrankTriplet}(r, X, Y, Z)$. $r$ is a rank and $X, Y, Z$ are the upper bounds for the numbers in the triplets.
Unranking Without Cross Products

UnrankingTreeNoCrossLocal(\(v, r, k\))

**Input:** an SDG node \(v\), a rank \(r\), a number \(k\) identifying a partition

**Output:** adornments of the SDG as a side-effect

Let \(\{w_1, \ldots, w_n\}\) be the children of \(v\)

```java
switch n {
    case 0:
        // no additional adornment for \(v\)
```
Unranking Without Cross Products (2)

\textbf{case 1:}

let \text{count}(v) = [c_0, \ldots, c_n]
let \text{summands}(v) = [s^0, \ldots, s^n]

k_1 = \min_j r \leq \sum_{i=0}^{j} s_i^k
r_1 = r - \sum_{i=0}^{k_1-1} s_i^k
insert-at(v) = k
\text{UnrankingTreeNoCrossLocal}(w_1, r_1, k_1)
Unranking Without Cross Products (3)

case 2:

let count(v) = [c_0, \ldots, c_n]
let summands(v) = [s^0, \ldots, s^n]
let count(w_1) = [c_0^1, \ldots, c_{n_1}^1]
let count(w_2) = [c_0^2, \ldots, c_{n_2}^2]

k_1 = \min_j r \leq \sum_{i=0}^{k_j} s_i^k
q = r - \sum_{i=0}^{k_1-1} s_i^k
k_2 = k - k_1

(r_1, r_2, a) = UnrankTriplet(q, c_{k_1}^1, c_{k_2}^2, \binom{k}{i})

\alpha = UnrankDecomposition(a)
merge-using(v) = \alpha

UnrankingTreeNoCrossLocal(w_1, r_1, k_1)
UnrankingTreeNoCrossLocal(w_2, r_2, k_2)
Quick Pick

- problem: build (pseudo-)random join trees fast
- unranking without cross products is quite involved
- idea: randomly select an edge in the query graph
- extend join tree by selected edge

No longer uniformly distributed, but very fast
Quick Pick (2)

QuickPick(Query Graph $G$)

**Input:** a query graph $G = (\{R_1, \ldots, R_n\}, E)$

**Output:** a bushy join tree

$E' = E$;

$\text{Trees} = \{R_1, \ldots, R_n\}$;

**while** $|\text{Trees}| > 1$ {

choose a random $e \in E'$

$E' = E' \setminus \{e\}$

if $e$ connects two relations in different subtrees $T_1, T_2 \in \text{Trees}$

$\text{Trees} = \text{Trees} \setminus \{T_1, T_2\} \cup \text{CreateJoinTree}(T_1, T_2)$

}

**return** $T \in \text{Trees}$

- repeated multiple times to find a good tree
Metaheuristics

- provide a very general optimization strategy
- applicable for many different problems
- work well even for very large problems
- but are often considered a "brute-force" method

We consider the metaheuristics formulated for the join ordering problem.
Iterative Improvement

- Start with random join tree
- Select rule that improves join tree
- Stop when no further improvement possible
Iterative Improvement (2)

IterativeImprovementBase(Query Graph $G$)

**Input:** a query graph $G = (\{ R_1, \ldots, R_n \}, E)$

**Output:** a join tree

```plaintext
do {
    JoinTree = random tree
    JoinTree = IterativeImprovement(JoinTree)
    if cost(JoinTree) < cost(BestTree) {
        BestTree = JoinTree
    }
} while (time limit not exceeded)
return BestTree
```
Iterative Improvement (3)

Iterative Improvement (Join Tree)

**Input:** a join tree

**Output:** improved join tree

```plaintext
do {
    JoinTree' = randomly apply a transformation from the rule set to the JoinTree
    if (cost(JoinTree') < cost(JoinTree)) {
        JoinTree = JoinTree'
    }
} while local minimum not reached
return JoinTree
```

- problem: local minimum detection
Simulated Annealing

- II: stuck in local minimum
- SA: allow moves that result in more expensive join trees
- lower the threshold for worsening
Simulated Annealing (2)

SimulatedAnnealing(Query Graph \( G \))

**Input:** a query graph \( G = (\{ R_1, \ldots, R_n \}, E) \)

**Output:** a join tree

BestTreeSoFar = random tree

Tree = BestTreeSoFar
Simulated Annealing (3)

\[
\text{do } \{ \\
\quad \text{do } \{ \\
\quad \quad \text{Tree' = apply random transformation to Tree} \\
\quad \quad \text{if (cost(Tree') < cost(Tree)) } \{ \\
\quad \quad \quad \text{Tree = Tree'} \\
\quad \quad \} \text{ else } \{ \\
\quad \quad \quad \text{with probability } e^{-(\text{cost(Tree') - cost(Tree)})/\text{temperature}} \\
\quad \quad \quad \quad \text{Tree = Tree'} \\
\quad \} \text{ if (cost(Tree) < cost(BestTreeSoFar)) } \{ \\
\quad \quad \quad \text{BestTreeSoFar = Tree'} \\
\quad \} \} \text{ while equilibrium not reached} \\
\quad \text{reduce temperature} \\
\} \text{ while not frozen} \\
\text{return BestTreeSoFar}
\]
Simulated Annealing (4)

Advantages:

- can escape from local minimum
- produces better results than II

Problems:

- parameter tuning
- initial temperature
- when and how to decrease the temperature
Tabu Search

- Select cheapest reachable neighbor (even if it is more expensive)
- Maintain tabu set to avoid running into circles
Tabu Search (2)

TabuSearch(Query Graph)

**Input:** a query graph \( G = (\{ R_1, \ldots, R_n \}, E) \)

**Output:** a join tree

Tree = random join tree

BestTreeSoFar = Tree

TabuSet = \( \emptyset \)

**do** {

Neighbors = all trees generated by applying a transformation to Tree

Tree = cheapest in Neighbors \ subtract TabuSet

if cost(Tree) < cost(BestTreeSoFar)

BestTreeSoFar = Tree

if (\( |\text{TabuSet}| > \text{limit} \)) remove oldest tree from TabuSet

TabuSet = TabuSet \cup \{ Tree \}

}

**return** BestTreeSoFar
Genetic Algorithms

- Join trees seen as population
- Successor generations generated by crossover and mutation
- Only the fittest survive

Problem: Encoding
- Chromosome $\leftrightarrow$ string
- Gene $\leftrightarrow$ character
Encoding

We distinguish \textit{ordered list} and \textit{ordinal number} encodings. Both encodings are used for left-deep and bushy trees. In all cases we assume that the relations $R_1, \ldots, R_n$ are to be joined and use the index $i$ to denote $R_i$. 
Ordered List Encoding

1. left-deep trees
   A left-deep join tree is encoded by a permutation of 1, \ldots, n. For instance, \(((R_1 \bowtie R_4) \bowtie R_2) \bowtie R_3\) is encoded as “1423”.

2. bushy trees
   A bushy join-tree without cartesian products is encoded as an ordered list of the edges in the join graph. Therefore, we number the edges in the join graph. Then, the join tree is encoded in a bottom-up, left-to-right manner.
Ordinal Number Encoding

In both cases, we start with the list $L = < R_1, \ldots, R_n >$.

- **left-deep trees**
  Within $L$ we find the index of first relation to be joined. If this relation be $R_i$ then the first character in the chromosome string is $i$. We eliminate $R_i$ from $L$. For every subsequent relation joined, we again determine its index in $L$, remove it from $L$ and append the index to the chromosome string.
  For instance, starting with $< R_1, R_2, R_3, R_4 >$, the left-deep join tree $(((R_1 \Join R_4) \Join R_2) \Join R_3)$ is encoded as “1311”.
Ordinal Number Encoding (2)

- bushy trees
  We encode a bushy join tree in a bottom-up, left-to-right manner. Let $R_i \bowtie R_j$ be the first join in the join tree under this ordering. Then we look up their positions in $L$ and add them to the encoding. Then we eliminate $R_i$ and $R_j$ from $L$ and push $R_{i,j}$ to the front of it. We then proceed for the other joins by again selecting the next join which now can be between relations and or subtrees. We determine their position within $L$, add these positions to the encoding, remove them from $L$, and insert a composite relation into $L$ such that the new composite relation directly follows those already present. For instance, starting with the list $< R_1, R_2, R_3, R_4 >$, the bushy join tree $((R_1 \bowtie R_2) \bowtie (R_3 \bowtie R_4))$ is encoded as “12 23 12”.
Crossover

1. Subsequence exchange
2. Subset exchange
Crossover: Subsequence exchange

The subsequence exchange for the ordered list encoding:

- Assume two individuals with chromosomes $u_1v_1w_1$ and $u_2v_2w_2$.
- From these we generate $u_1v_1'w_1$ and $u_2v_2'w_2$ where $v_i'$ is a permutation of the relations in $v_i$ such that the order of their appearance is the same as in $u_3−i v_3−i w_3−i$.

Subsequence exchange for ordinal number encoding:

- We require that the $v_i$ are of equal length ($|v_1| = |v_2|$) and occur at the same offset ($|u_1| = |u_2|$).
- We then simply swap the $v_i$.
- That is, we generate $u_1v_2w_1$ and $u_2v_1w_2$. 
Crossover: Subset exchange

The subset exchange is defined only for the ordered list encoding. Within the two chromosomes, we find two subsequences of equal length comprising the same set of relations. These sequences are then simply exchanged.
Mutation

A mutation randomly alters a character in the encoding. If duplicates may not occur—as in the ordered list encoding—swapping two characters is a perfect mutation.
Selection

• The probability of survival is determined by its rank in the population.
• We calculate the costs of the join trees encoded for each member in the population.
• Then, we sort the population according to their associated costs and assign probabilities to each individual such that the best solution in the population has the highest probability to survive and so on.
• After probabilities have been assigned, we randomly select members of the population taking into account these probabilities.
• That is, the higher the probability of a member the higher its chance to survive.
The Algorithm

1. Create a random population of a given size (say 128).
2. Apply crossover and mutation with a given rate. For example such that 65% of all members of a population participate in crossover, and 5% of all members of a population are subject to random mutation.
3. Apply selection until we again have a population of the given size.
4. Stop after no improvement within the population was seen for a fixed number of iterations (say 30).
Combinations

- metaheuristics are often not used in isolation
- they can be used to improve existing heuristics
- or heuristics can be used to speed up metaheuristics
Two Phase Optimization

1. For a number of randomly generated initial trees, Iterative Improvement is used to find a local minima.

2. Then Simulated Annealing is started to find a better plan in the neighborhood of the local minima. The initial temperature of Simulated Annealing can be lower as is its original variants.
AB Algorithm

1. If the query graph is cyclic, a spanning tree is selected.
2. Assign join methods randomly
3. Apply IKKBZ
4. Apply iterative improvement
Toured Simulated Annealing

The basic idea is that simulated annealing is called $n$ times with different initial join trees, if $n$ is the number of relations to be joined.

- Each join sequence in the set $S$ produced by GreedyJoinOrdering-3 is used to start an independent run of simulated annealing.

As a result, the starting temperature can be decreased to 0.1 times the cost of the initial plan.
GOO-II

Append an iterative improvement step to GOO
Iterative Dynamic Programming

- Two variants: IDP-1, IDP-2 [9]
- Here: Only IDP-1 base version

Idea:
- create join trees with up to $k$ relations
- replace cheapest one by a compound relation
- start all over again
Iterative Dynamic Programming (2)

IDP-1({R_1, \ldots, R_n}, k)
Input: a set of relations to be joined, maximum block size k
Output: a join tree
for each 1 \leq i \leq n {
    BestTree({R_i}) = R_i;
}
ToDo = \{R_1, \ldots, R_n\}
Iterative Dynamic Programming (3)

while |ToDo| > 1 {
    k = min(k, |ToDo|)
    for each 2 ≤ i < k ascending
        for all S ⊆ ToDo, |S| = i do
            for all O ⊂ S do
                BestTree(S) = CreateJoinTree(BestTree(S \ O), BestTree(O));
        find V ⊂ ToDo, |V| = k with
            cost(BestTree(V)) = min{cost(BestTree(W)) | W ⊂ ToDo, |W| = k}
        generate new symbol T
        BestTree({T}) = BestTree(V)
        ToDo = (ToDo \ V) ∪ {T}
    for each O ⊂ V do delete(BestTree(O))
}
return BestTree({R₁, . . . , Rₙ})
Iterative Dynamic Programming (4)

- compromise between runtime and optimality
- combines greedy heuristics with dynamic programming
- scales well to large problems
- finds the optimal solution for smaller problems
- approach can be used for different DP strategies
Order Preserving Joins

- some query languages operate on lists instead of sets/bags
- order of tuples matters
- examples: XPath/XQuery
- alternatives: either add sort operators or use order preserving operators

Here, we define order preserving operators, \( \text{list} \rightarrow \text{list} \)

- let \( L \) be a list
- \( L[1] \) is the first entry in \( L \)
- \( L[2 : |L|] \) are the remaining entries
Order Preserving Selection

We define the order preserving selection $\sigma^L$ as follows:

$$
\sigma^L_p(e) := \begin{cases} 
\epsilon & \text{if } e = \epsilon \\
< e[1] > \circ \sigma^L_p(e[2 : |e|]) & \text{if } p(e[1]) \\
\sigma^L_p(e[2 : |e|]) & \text{otherwise}
\end{cases}
$$

- filters like a normal selection
- preserves the relative ordering (guaranteed)
Order Preserving Cross Product

We define the order preserving cross product $\times^L$ as follows:

$$e_1 \times^L e_2 := \begin{cases} 
\epsilon & \text{if } e_1 = \epsilon \\
(e[1] \hat{\times}^L e_2) \circ (e_1[2 : \mid e_1 \mid] \times^L e_2) & \text{otherwise}
\end{cases}$$

using the tuple/list product defined as:

$$t \hat{\times}^L e := \begin{cases} 
\epsilon & \text{if } e = \epsilon \\
< t \circ e[1] > \circ (t \hat{\times}^L e[2 : \mid e \mid]) & \text{otherwise}
\end{cases}$$

- preserves the order of $e_1$
- order of $e_2$ is preserved for each $e_1$ group
Order Preserving Join

The definition of the order preserving join is analogous to the non-order preserving case:

\[ e_1 \bowtie_p^L e_2 := \sigma_p^L (e_1 \times^L e_2) \]

- preserves order of \( e_1 \), order of \( e_2 \) relative to \( e_1 \)
Equivalences

\[ \sigma_{p_1}^L(\sigma_{p_2}^L(e)) \equiv \sigma_{p_2}^L(\sigma_{p_1}^L(e)) \]
\[ \sigma_{p_1}^L(e_1 \bowtie_{p_2}^L e_2) \equiv \sigma_{p_1}^L(e_1) \bowtie_{p_2}^L e_2 \quad \text{if } \mathcal{F}(p_1) \subseteq \mathcal{A}(e_1) \]
\[ \sigma_{p_2}^L(e_1 \bowtie_{p_2}^L e_2) \equiv e_1 \bowtie_{p_2}^L \sigma_{p_1}^L(e_2) \quad \text{if } \mathcal{F}(p_1) \subseteq \mathcal{A}(e_2) \]
\[ e_1 \bowtie_{p_1}^L (e_2 \bowtie_{p_2}^L e_3) \equiv (e_1 \bowtie_{p_1}^L e_2) \bowtie_{p_2}^L e_3 \quad \text{if } \mathcal{F}(p_i) \subseteq \mathcal{A}(e_i) \cup \mathcal{A}(e_{i+1}) \]

- swap selections
- push selections down
- associativity
Commutativity

Consider the relations $R_1 = \langle [a : 1], [a : 2] \rangle$ and $R_2 = \langle [b : 1], [b : 2] \rangle$. Then

$$R_1 \bowtie^L_{true} R_2 = \langle [a : 1, b : 1], [a : 1, b : 2], [a : 2, b : 1], [a : 2, b : 2] \rangle$$
$$R_2 \bowtie^L_{true} R_1 = \langle [a : 1, b : 1], [a : 2, b : 1], [a : 1, b : 2], [a : 2, b : 2] \rangle$$

- the order preserving join is not commutative
Algorithm

- similar to matrix multiplication
- in addition: selection push down
- DP table is a $n \times n$ array (or rather 4 arrays)
- algorithm fills arrays $p, s, c, t$:
  - $p$: applicable predicates
  - $s$: statistics (cardinality, perhaps more)
  - $c$: costs
  - $t$: split position for larger plans
- plan is extracted from the arrays afterwards
Algorithm (2)

OrderPreservingJoins($R = \{ R_1, \ldots, R_n \}, P$)

**Input:** a set of relations to be joined and a set of predicates

**Output:** fills $p, s, c, t$

**for each** $1 \leq i \leq n$

1. $p[i, i] =$predicates from $P$ applicable to $R_i$
2. $P = P \setminus p[i, i]$
3. $s[i, i] =$statistics for $\sigma_{p[i, i]}(R_i)$
4. $c[i, i] =$costs for $\sigma_{p[i, i]}(R_i)$
Algorithm (3)

for each $2 \leq l \leq n$ ascending {
    for each $1 \leq i \leq n - l + 1$ {
        $j = i + l - 1$
        $p[i, j] =$ predicates from $P$ applicable to $R_i, \ldots, R_j$
        $P = P \setminus p[i, j]$
        $s[i, j] =$ statistics derived from $s[i, j - 1]$ and $s[j, j]$ including $p[i, j]$
        $c[i, j] = \infty$
        for each $i \leq k < j$ {
            $q = c[i, k] + c[k + 1, j] + \text{costs for } s[i, k] \text{ and } s[k + 1, j] \text{ and } p[i, j]$
            if $q < c[i, j]$ {
                $c[i, j] = q$
                $t[i, j] = k$
            }
        }
    }
}
Algorithm (4)

ExtractPlan\((R = \{R_1, \ldots, R_n\}, t, p)\)

**Input:** a set of relations, arrays \(t\) and \(p\)

**Output:** a bushy join tree

return ExtractPlanRec\((R, t, p, 1, n)\)

ExtractPlanRec\((R = \{R_1, \ldots, R_n\}, t, p, i, j)\)

if \(i < j\) {
   \(T_1 = \text{ExtractPlanRec}(R, t, p, i, t[i, j])\)
   \(T_2 = \text{ExtractPlanRec}(R, t, p, t[i, j] + 1, j)\)
   return \(T_1 \Join^L_p[i, j] T_2\)
} else {
   return \(\sigma_p[i, j] R_i\)
}