Query Optimization

Exercise Session 7

Andrey Gubichev

June 2, 2014

DPccp

- Enumerate over all connected subgraphs
- For each subgraph enumerate all other connected subgraphs that are disjoint but connected to it

Enumerating Complementary Subgraphs

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, $(S \cup S')$, $(X \cup N)$);
}

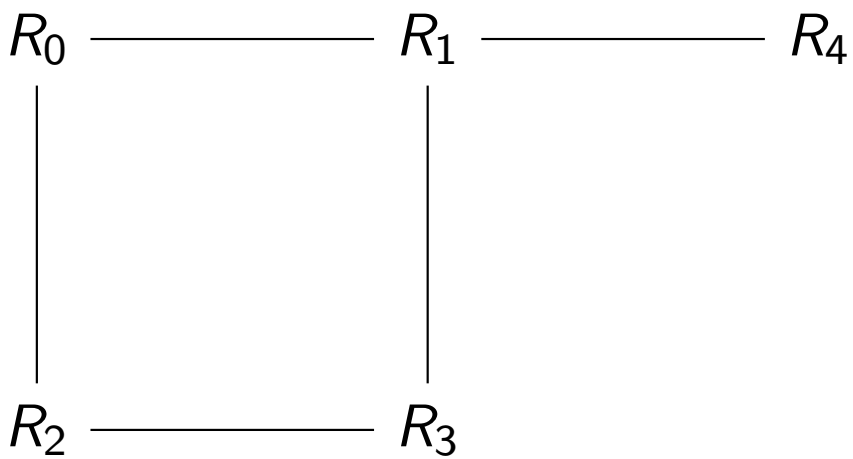
Enumerating Complementary Subgraphs

EnumerateCmp(G,S_1)
$X = \mathcal{B}_{\min(S_1)} \cup S_1$;
$N = \mathcal{N}(S_1) \setminus X$;
for all ($v_i \in N$ by descending i) {
    emit $\{v_i\}$;
    EnumerateCsgRec(G, $\{v_i\}$, $X \cup (\mathcal{B}_i \cap N)$);
}

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

Homework: Task 1

Given the following query graph, enumerate all connected subgraph-complement-pairs as produced by DPccp (not just connected subgraphs!):

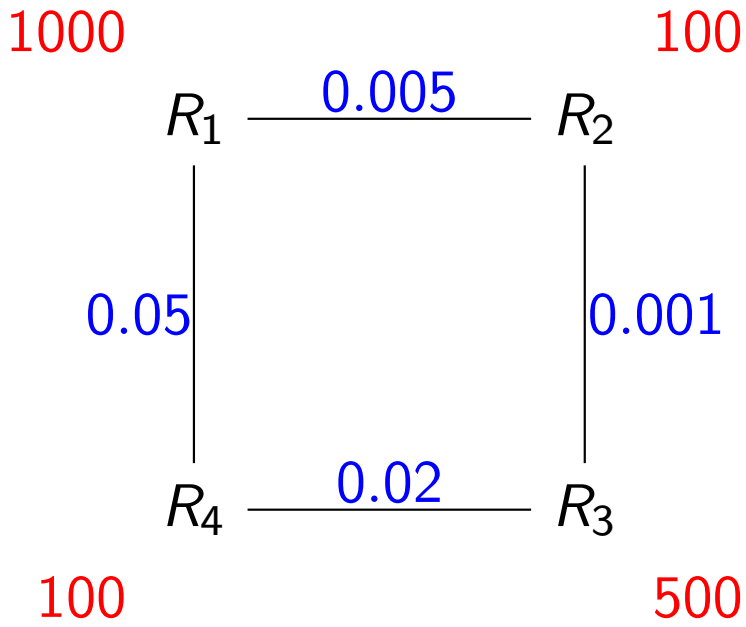


Graph simplification

Sometimes the graph is too big, let's simplify it.

- GOO: choose the joins greedily (very hard, depends on all other joins)
- Simplification: choose the joins that must be avoided (we can start with 'obvious' decisions)

Graph simplification: Example



- $benefit(X \bowtie R_1, X \bowtie R_2) = \frac{C((X \bowtie R_1) \bowtie R_2)}{C((X \bowtie R_2) \bowtie R_1)}$
- $R_3 \bowtie R_2$ before $R_3 \bowtie R_4$. Remove $R_4 - R_3$
- $R_4 \bowtie (R_2 \bowtie R_3)$ before $R_4 \bowtie R_1$. Remove $R_1 - R_4$
- no more choices

More insights

- Guido Moerkotte, Thomas Neumann. Dynamic Programming Strikes Back. In *SIGMOD'08*
- Thomas Neumann. Query Simplification: Graceful Degradation for Join-Order Optimization. In *SIGMOD'09*

Info

- Exercises due: 9 AM, June 9, 2014