

Exercises for the book

Semantic Web Databases

by Sven Groppe

Chapter 5 – Logical Optimization

1. Transform the following SPARQL query into an algebra expression using the algebra as proposed in this chapter:

```

PREFIX rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf:   <http://xmlns.com/foaf/0.1/>
PREFIX bench: <http://localhost/vocabulary/bench/>
PREFIX dc:     <http://purl.org/dc/elements/1.1/>
SELECT DISTINCT ?name
WHERE {
  ?article rdf:type bench:Article .
  ?article dc:creator ?person .
  ?inproc dc:creator ?person .
  ?person foaf:name ?name
} ORDER BY ?name

```

2. Optimize the following algebraic expression according to the heuristic approach using equivalency rules:

```

(((?article dc:creator ?person) AND
  (?person2 foaf:name ?name))
  FILTER(?name <= "M"))
  FILTER(?person2 = ?person)

```

3. Use dynamic programming to optimize the join order of a join between the triple patterns $t_1 = (?article \text{ rdf:type } bench:Article)$, $t_2 = (?article \text{ dc:creator } ?person)$, $t_3 = (?inproc \text{ dc:creator } ?person)$ and $t_4 = (?person \text{ foaf:name } ?name)$. Do not estimate the costs but calculate the number of all intermediate results for the costs. The following table already displays the results of the triple patterns and their join partners.

t_1	t_2		t_3		t_4	
<i>?article</i>	<i>?article</i>	<i>?person</i>	<i>?person</i>	<i>?inproc</i>	<i>?person</i>	<i>?name</i>
<a1>	<a1>	<p1>	<p1>	<i1>	<p1>	<n1>
	<a2>	<p2>	<p2>	<i2>		
	<a3>	<p3>			<p3>	<n1>
	<a4>	<p3>				