Cs5356:Implementation of the SDD-1 Algorithm
Jialin Liu. Mar 31.2012

1. Description

This is the implementation of the SDD-1 Algorithm as described in the Ozsu and Valduriez textbook (Algorithm 8.6). Program design and output display are included in this report.

2. Program

Programming language: C
Compile: gcc 4.0
Runtime environment: Linux 2.6
Compile command: gcc –o sdd sdd.c
Running command: ./sdd
This program is designed to be general purpose, which means the program is not only specified to this one example.

3. Input format

In this example, we have the following information
Relations

Name	Size
R1	1000
R2	1000
R3	2000
R4	1000

Connections

Relation1	Relation2	Attribute
R1	R2	A
R2	R3	B
R2	R4	B
R3	R4	b

Statistics

Relation	Attribute	Size	SFsj
R1	A	200	0.5
R2	A	100	0.1
R2	B	100	0.2
R3	B	300	0.9
R4	B	150	0.4

When running the program, users are required to input as the following format:
please input the number of relations:
4
please input the relation size
R1:1000
R2:1000
R3:2000
R4:1000

please input the message cost:
20
please input the transmission cost:
1
please input the number of connections between relations:
4
input the 1 connection:(eg.1 2 A)
1 2 a
input the 2 connection:(eg.1 2 A)
2 3 b
input the 3 connection:(eg.1 2 A)
2 4 b
input the 4 connection:(eg.1 2 A)
3 4 b
please input the number of statistics tuples
5
please input the initial projection statistics
e.g:1 36 0.3 A

relation	size	selection_factor	attribute
1 200 0.5 a			

relation	size	selection_factor	attribute
2 100 0.1 a			

relation	size	selection_factor	attribute
2 100 0.2 b			

relation	size	selection_factor	attribute
3 300 0.9 b			

relation	size	selection_factor	attribute
4 150 0.4 b			

4. Output display

After inputting the above information in the terminal. The program will run automatically to the final loop, the result of each loop will be displayed.

```
*************************************************
```

Initial status:

```
*************************************************

the list of possible semijoins with their cost and benefit:
semijoin,        benefit,          cost
r1(r1.a) SJ r2(r2.a),       900.000000,       120.000000
r2(r2.b) SJ r3(r3.b),       100.000023,       320.000000
r2(r2.b) SJ r4(r4.b),       600.000000,       170.000000
r3(r3.b) SJ r4(r4.b),      1200.000000,       170.000000
r2(r2.a) SJ r1(r1.a),       500.000000,       220.000000
r3(r3.b) SJ r2(r2.b),      1600.000000,       120.000000
r4(r4.b) SJ r2(r2.b),       800.000000,       120.000000
r4(r4.b) SJ r3(r3.b),       100.000023,       320.000000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,    SFsj
r1.a, 200.000000,      0.500000
r2.a, 100.000000,      0.100000
r2.b, 100.000000,      0.200000
r3.b, 300.000000,      0.900000
r4.b, 150.000000,      0.400000
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,        size
r1,      1000.000000
r2,      1000.000000
r3,      2000.000000
r4,      1000.000000


*************************************************
```

1st loop

```
*************************************************
the list of possible semijoins with their cost and benefit:
semijoin,        benefit,          cost
r1(r1.a) SJ r2(r2.a),       900.000000,       120.000000
r2(r2.b) SJ r3(r3.b),       820.000000,        80.000000
r2(r2.b) SJ r4(r4.b),       600.000000,       170.000000
r3(r3.b) SJ r4(r4.b),       240.000015,       170.000000
r2(r2.a) SJ r1(r1.a),       500.000000,       220.000000
r4(r4.b) SJ r2(r2.b),       800.000000,       120.000000
r4(r4.b) SJ r3(r3.b),       820.000000,        80.000000
the most beneficial semijoins with their cost and benefit:
semijoin,        benefit,          cost
r3(r3.b) SJ r2(r2.b),      1600.000000,       120.000000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,    SFsj
r1.a, 200.000000,      0.500000
r2.a, 100.000000,      0.100000
r2.b, 100.000000,      0.200000
r3.b, 60.000000,      0.180000
r4.b, 150.000000,      0.400000
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,        size
r1,      1000.000000
r2,      1000.000000
r3,      400.000000
r4,      1000.000000
```

```
**************************************************
2nd loop
**************************************************
the list of possible semijoins with their cost and benefit:
semijoin,         benefit,           cost
r2(r2.b) SJ r3(r3.b),      820.000000,       80.000000
r2(r2.b) SJ r4(r4.b),      600.000000,      170.000000
r3(r3.b) SJ r4(r4.b),      240.000015,      170.000000
r2(r2.a) SJ r1(r1.a),      950.000000,       40.000000
r4(r4.b) SJ r2(r2.b),      800.000000,      120.000000
r4(r4.b) SJ r3(r3.b),      820.000000,       80.000000
the most beneficial semijoins with their cost and benefit:
semijoin,         benefit,           cost
r3(r3.b) SJ r2(r2.b),     1600.000000,      120.000000
r1(r1.a) SJ r2(r2.a),      900.000000,      120.000000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,   SFsj
r1.a, 20.000000,     0.050000
r2.a, 100.000000,     0.100000
r2.b, 100.000000,     0.200000
r3.b, 60.000000,     0.180000
r4.b, 150.000000,     0.400000
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,        size
r1,     100.000000
r2,     1000.000000
r3,     400.000000
r4,     1000.000000
**************************************************
3rd loop
**************************************************
the list of possible semijoins with their cost and benefit:
semijoin,         benefit,           cost
r2(r2.b) SJ r3(r3.b),       41.000000,       80.000000
r2(r2.b) SJ r4(r4.b),       30.000002,      170.000000
r3(r3.b) SJ r4(r4.b),      240.000015,      170.000000
r4(r4.b) SJ r2(r2.b),      900.000000,       70.000000
r4(r4.b) SJ r3(r3.b),      820.000000,       80.000000
the most beneficial semijoins with their cost and benefit:
semijoin,         benefit,           cost
r3(r3.b) SJ r2(r2.b),     1600.000000,      120.000000
r1(r1.a) SJ r2(r2.a),      900.000000,      120.000000
r2(r2.a) SJ r1(r1.a),      950.000000,       40.000000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,   SFsj
r1.a, 20.000000,     0.050000
r2.a, 5.000000,     0.005000
r2.b, 50.000000,     0.100000
r3.b, 60.000000,     0.180000
r4.b, 150.000000,     0.400000
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,        size
r1,     100.000000
r2,     50.000000
r3,     400.000000
r4,     1000.000000
```

```
**************************************************
```

4th loop

```
**************************************************
the list of possible semijoins with their cost and benefit:
semijoin,        benefit,         cost
r2(r2.b) SJ r3(r3.b),     41.000000,      80.000000
r2(r2.b) SJ r4(r4.b),     48.000000,      35.000000
r3(r3.b) SJ r4(r4.b),     384.000000,      35.000000
r4(r4.b) SJ r3(r3.b),     82.000000,      80.000000
the most beneficial semijoins with their cost and benefit:
semijoin,        benefit,         cost
r3(r3.b) SJ r2(r2.b),     1600.000000,     120.000000
r1(r1.a) SJ r2(r2.a),     900.000000,     120.000000
r2(r2.a) SJ r1(r1.a),     950.000000,      40.000000
r4(r4.b) SJ r2(r2.b),     900.000000,      70.000000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,    SFsj
r1.a, 20.000000,      0.050000
r2.a, 5.000000,       0.005000
r2.b, 50.000000,      0.100000
r3.b, 60.000000,      0.180000
r4.b, 15.000000,      0.040000
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,        size
r1,      100.000000
r2,      50.000000
r3,      400.000000
r4,      100.000000
**************************************************
```

5th loop

```
**************************************************
the list of possible semijoins with their cost and benefit:
semijoin,        benefit,         cost
r2(r2.b) SJ r3(r3.b),     49.639999,      22.400000
r2(r2.b) SJ r4(r4.b),     48.000000,      35.000000
r4(r4.b) SJ r3(r3.b),     99.279999,      22.400000
the most beneficial semijoins with their cost and benefit:
semijoin,        benefit,         cost
r3(r3.b) SJ r2(r2.b),     1600.000000,      120.000000
r1(r1.a) SJ r2(r2.a),     900.000000,      120.000000
r2(r2.a) SJ r1(r1.a),     950.000000,      40.000000
r4(r4.b) SJ r2(r2.b),     900.000000,      70.000000
r3(r3.b) SJ r4(r4.b),     384.000000,      35.000000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,    SFsj
r1.a, 20.000000,      0.050000
r2.a, 5.000000,       0.005000
r2.b, 50.000000,      0.100000
r3.b, 2.400000,       0.007200
r4.b, 15.000000,      0.040000
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,        size
r1,      100.000000
r2,      50.000000
r3,      16.000002
r4,      100.000000
```

```
*************************************************
```
6th loop
```
*************************************************
the list of possible semijoins with their cost and benefit:
semijoin,          benefit,          cost
r2(r2.b) SJ r3(r3.b),      49.639999,      22.400000
r2(r2.b) SJ r4(r4.b),      49.985600,      20.108000
the most beneficial semijoins with their cost and benefit:
semijoin,          benefit,          cost
r3(r3.b) SJ r2(r2.b),      1600.000000,      120.000000
r1(r1.a) SJ r2(r2.a),       900.000000,      120.000000
r2(r2.a) SJ r1(r1.a),       950.000000,       40.000000
r4(r4.b) SJ r2(r2.b),       900.000000,       70.000000
r3(r3.b) SJ r4(r4.b),       384.000000,       35.000000
r4(r4.b) SJ r3(r3.b),        99.279999,      22.400000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,    SFsj
r1.a, 20.000000,      0.050000
r2.a, 5.000000,       0.005000
r2.b, 50.000000,       0.100000
r3.b, 2.400000,       0.007200
r4.b, 0.108000,       0.000288
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,          size
r1,      100.000000
r2,       50.000000
r3,       16.000002
r4,        0.720000
*************************************************
```
Final loop
```
*************************************************
the list of possible semijoins with their cost and benefit:
semijoin,          benefit,          cost
r2(r2.b) SJ r3(r3.b),       0.014296,      22.400000
the most beneficial semijoins with their cost and benefit:
semijoin,          benefit,          cost
r3(r3.b) SJ r2(r2.b),      1600.000000,      120.000000
r1(r1.a) SJ r2(r2.a),       900.000000,      120.000000
r2(r2.a) SJ r1(r1.a),       950.000000,       40.000000
r4(r4.b) SJ r2(r2.b),       900.000000,       70.000000
r3(r3.b) SJ r4(r4.b),       384.000000,       35.000000
r4(r4.b) SJ r3(r3.b),        99.279999,      22.400000
r2(r2.b) SJ r4(r4.b),        49.985600,      20.108000
the changes of statistics caused by execution of the most beneficial semijoin:
projection,      size,    SFsj
r1.a, 20.000000,      0.050000
r2.a, 0.014400,       0.000014
r2.b, 0.014400,       0.000029
r3.b, 2.400000,       0.007200
r4.b, 0.108000,       0.000288
the changes of statistics of realtions caused by execution of the most beneficial semijoin
relation,          size
r1,      100.000000
r2,        0.014400
r3,       16.000002
r4,        0.720000
```
Assembly site

```
the assembly size for the query is:
 r1
```
/*A post optimization has not been implemented yet.*/