CS-410: Database Engineering
Medium-stakes Assignment: Relational Algebra Queries

Nick Alexander

February 13, 2015

Contents

1 Relational Algebra Queries **2**

2 Rubric **9**

3 Self-assessment **10**

1 Relational Algebra Queries

1. List student ids and last names for students whose first name is 'Joseph'

   $\sigma_{fname='Joseph'}$(Student)

sid	fname	lname
3	Joseph	Green

2. List course id, course name, credit hours, and prerequisite course number for all courses with credit hours greater than 3.0

   $\sigma_{hours>3}$(Course)

cid	cname	hours	prereq
math229	Calculus	5.0	math130
math329	Linear Algebra	4.0	math229
cs110	Computer Science I	4.0	math1430
cs305	Software Engineering	4.0	math330

3. For each course, list its course name and prerequisite number.

   $\pi_{cname,prereq}$(Course)

College Algebra	NULL
Calculus I	math130
Discrete Structures	math229
Linear Algebra	math229
Computer Science I	math130
Computer Science II	cs110
Data Structures	math220
Programming Languages	cs210
Software Engineering I	math220
Database Engineering	math229
Image Processing	math329

4. List the names of all professional organizations.

   π_{oname}(Org)

oname
IEEE Computer Society
Computer Society of Indea
ACM

5. List all possible combinations of Student and Org tuples.

   Student $\times$ Org

sid	fname	lname	oid	oname	fee
1	William	Campbell	1	IEEE Computer Society	45.00
1	William	Campbell	2	Computer Society of India	25.00
1	William	Campbell	3	ACM	55.00
2	Robert	Hill	1	IEEE Computer Society	45.00
2	Robert	Hill	2	Computer Society of India	25.00
2	Robert	Hill	3	ACM	55.00
3	Joseph	Green	1	IEEE Computer Society	45.00
3	Joseph	Green	2	Computer Society of India	25.00
3	Joseph	Green	3	ACM	55.00
4	Jeff	Wilson	1	IEEE Computer Society	45.00
4	Jeff	Wilson	2	Computer Society of India	25.00
4	Jeff	Wilson	3	ACM	55.00
5	Patricia	Davis	1	IEEE Computer Society	45.00
5	Patricia	Davis	2	Computer Society of India	25.00
5	Patricia	Davis	3	ACM	55.00
6	Susan	Brown	1	IEEE Computer Society	45.00
6	Susan	Brown	2	Computer Society of India	25.00
6	Susan	Brown	3	ACM	55.00
7	Thomas	Smith	1	IEEE Computer Society	45.00
7	Thomas	Smith	2	Computer Society of India	25.00
7	Thomas	Smith	3	ACM	55.00
8	Mark	Williams	1	IEEE Computer Society	45.00
8	Mark	Williams	2	Computer Society of India	25.00
8	Mark	Williams	3	ACM	55.00
9	Paul	Jones	1	IEEE Computer Society	45.00
9	Paul	Jones	2	Computer Society of India	25.00
9	Paul	Jones	3	ACM	55.00
10	Barbara	Robinson	1	IEEE Computer Society	45.00
10	Barbara	Robinson	2	Computer Society of India	25.00
10	Barbara	Robinson	3	ACM	55.00
11	Jennifer	King	1	IEEE Computer Society	45.00
11	Jennifer	King	2	Computer Society of India	25.00
11	Jennifer	King	3	ACM	55.00
12	Sarah	Parker	1	IEEE Computer Society	45.00
12	Sarah	Parker	2	Computer Society of India	25.00
12	Sarah	Parker	3	ACM	55.00
13	Lisa	Lopez	1	IEEE Computer Society	45.00
13	Lisa	Lopez	2	Computer Society of India	25.00
13	Lisa	Lopez	3	ACM	55.00
14	Sharon	Jackson	1	IEEE Computer Society	45.00
14	Sharon	Jackson	2	Computer Society of India	25.00
14	Sharon	Jackson	3	ACM	55.00
15	Kevin	Miller	1	IEEE Computer Society	45.00
15	Kevin	Miller	2	Computer Society of India	25.00
15	Kevin	Miller	3	ACM	55.00

6. List the names of attributes in the result set produced by executing query 5?

   The attributes in the new relation from query 5 are sid, fname, lname, oid, oname, and fee

7. What is the cardinality of the result set produced by executing query 5? In other words, how may tuples are there in the result set?

   The cardinality of the new relation from query 5 is 45.

8. Given the cardinalities of Student and Org relations, how do you compute the number of tuples in the CARTESIAN PRODUCT of Student and Org?

   In order to calculate the cardinality of a resulting relation from a Cartesian product operation, one must multiply the cardinality of relation R by the cardinality of relation S. In query 5, the relations R and S are Student and Org.

9. For students enrolled in courses, list student id, first name, last name, section id, and grade earned.

   Student ⋈ Enrollment

sid	fname	lname	secid	grade
1	William	Campbell	1	C
2	Robert	Hill	1	B
5	Patricia	Davis	6	B
5	Patricia	Davis	5	A
6	Susan	Brown	10	F
7	Thomas	Smith	4	B
8	Mark	Williams	6	A
8	Mark	Williams	5	A
9	Paul	Jones	9	D
15	Kevin	Miller	10	C

10. Examine the result set produced by executing query 9. Explain how the tuples in the result set are formed.

    Because sid is similar between both tables, the tuples are joined when Student sid matches Enrollment sid.

11. For students who have membership in professional organizations, list student id, first name, last name, organization id, membership start date, and membership end date (if any).

    Student ⋈ Membership

sid	fname	lname	oid	sdate	edate
1	William	Campbell	1	2006-01-01	NULL
1	William	Campbell	2	2009-02-08	NULL
1	William	Camp ell	3	2009-05-06	NULL
2	Robert	Hill	1	2007-05-01	NULL
5	Patricia	Davis	1	2008-09-08	NULL
5	Patricia	Davis	2	2009-06-02	NULL
5	Patricia	Davis	3	2007-11-30	NULL
6	Susan	Brown	1	2006-12-05	2009-09-10
8	Mark	Williams	2	2008-08-15	2009-08-14
15	Kevin	Miller	3	2005-06-03	2010-01-05

12. For query 11, list only the student's first name, last name, organization id, and membership start date.

    $\pi_{fname,lname,oid,sdate}(\text{Student} \bowtie \text{Membership})$

fname	lname	oid	sdate
William	Campbell	1	2006-01-01
William	Campbell	2	2009-02-08
William	Camp ell	3	2009-05-06
Robert	Hill	1	2007-05-01
Patricia	Davis	1	2008-09-08
Patricia	Davis	2	2009-06-02
Patricia	Davis	3	2007-11-30
Susan	Brown	1	2006-12-05
Mark	Williams	2	2008-08-15
Kevin	Miller	3	2005-06-03

13. List id numbers of students who are either members of organizations or enrolled in some section of a course(s).

    In order to satisfy union compatibility, we must project only the sid from enrollment and membership in order to apply the union operator.

    $\pi_{sid}(\pi_{sid}(\text{Enrollment}) \cup \pi_{sid}(\text{Membership}))$

sid
1
2
5
6
8
15
7
9

14. List student id numbers, first names, and last names of students who are either members of organizations or enrolled in some (section of a) course(s). Recognize that this query requires only a slight modification of query 13.

    $\pi_{sid,fname,lname}((\pi_{sid}(\text{Enrollment}) \cup \pi_{sid}(\text{Membership})) \bowtie \text{Student})$

sid	fname	lname
1	William	Campbell
2	Robert	Hill
5	Patricia	Davis
6	Susan	Brown
8	Mark	Williams
15	Kevin	Miller
7	Thomas	Smith
9	Paul	Jones

15. List the id numbers of students who are members of organizations as well as enrolled in some section of a course.

    Like before, we must project the sid of each relation in order to preserve union compatibility.

    $\pi_{sid}(\pi_{sid}(\text{Enrollment}) \cap \pi_{sid}(\text{Membership}))$

sid
1
2
5
6
8
15

16. Retrieve the ids of students who are members of organizations, but are not enrolled in sections of courses.

    $\pi_{sid}(\pi_{sid}(\text{Membership}) \setminus \pi_{sid}(\text{Enrollment}))$

sid

17. List the ids of students who are members of every organization.

    temp1 $\leftarrow \pi_{sid,oid}(\text{Membership})$
    temp2 $\leftarrow \pi_{oid}(\text{Org})$
    result $\leftarrow$ temp1 $\div$ temp2

sid
1
5

18. List the ids of all students who have enrolled in every section that the student with id number 5 has enrolled in. Hint: Write the query using the DIVISION operator and manually execute it. Show your result set.

    temp1 $\leftarrow \pi_{secid}(\sigma_{sid=5}(\text{Enrollment}))$
    result $\leftarrow \pi_{sid}(\text{Enrollment} \div \text{temp1})$

sid
5
8

19. Create a copy of relation Org and rename the attributes (oid, oname, fee) as (orgid, orgname, dues).

    $\rho_{orgid,orgname,dues}(\text{Org})$

orgid	orgname	dues
1	IEEE Computer Society	45.00
2	Computer Society of India	25.00
3	ACM	55.00

2 Rubric

Use the following rubric to evaluate your response to this assignment.

Perf Level / *Trait*	*Outstanding*	*Good*	*Fair*	*Poor*
Problem Analysis	Precise and concise documentation to support the claim that the problem is correctly and thoroughly analyzed prior to writing the query is provided.	Sufficient documentation, though verbose and non-coherent, to support the claim that the problem is correctly analyzed prior to writing the query is provided.	Some documentation to support the claim that the problem is superficially analyzed prior to writing the query is provided.	Documentation to support the claim that the problem is analyzed prior to writing the query is not provided.
Query Execution	The query compiles and produces correct results. Also, the query runs efficiently.	Though inefficient, the query compiles and produces correct results.	The query compiles but produces incorrect results.	The query does not compile.
Correctness Arguments	The correctness of the query is argued using rigorous statements.	The correctness of the query is argued using informal logical statements.	The correctness of the query is argued using illogical statements.	There are no statement about the query correctness.
Completeness	Answers to all questions in the assignment are provided. Queries compile and run. Problem analysis documentation and correctness arguments are provided.	Answers to less than 75% of the questions in the assignment are provided. Queries compile and run. Problem analysis documentation and correctness arguments are provided.	Answers to less than 50% of the questions in the assignment are provided. Queries compile and run. Problem analysis documentation and correctness arguments are provided.	Answers to less than 25% of the questions in the assignment are provided. Queries compile and run. Problem analysis documentation and correctness arguments are provided.

3 Self-assessment

Use the following table and the rubric of section 2 to score your solution. Circle the appropriate number in each row. For example, to circle 20, use the LaTeX markup code `\circled{20}`, which produces (20).

Perf Level / *Trait*	*Outstanding*	*Good*	*Fair*	*Poor*
Problem Analysis	(10)	8	6	4
Query Execution	(30)	25	20	15
Correctness Arguments	(10)	8	6	4
Completeness	(50)	40	30	20