**CSCI 4333.2**

10/14/2024

**Introduction to Relational Algebra and Relational Calculus**

by K. Yue

**1. Introduction**

* More theoretical query languages of the relational model:
	1. Relational Algebra (RA): a simple functional language; function calls (LISP)
	2. Relational Calculus (RC): a simple declarative language

Programming paradigms:

* Procedural: steps by steps; how?
* Declarative: what?
* Provide theoretical foundation for the relational model.
* Not used in practical DBMS.
* Some symbols in RA/RC: [RA\_RC\_Symbols.docx](https://dcm.uhcl.edu/yue/courses/joinDB/Fall2024/notes/rarc/RA_RC_Symbols.docx)

**2. Introduction to Relational algebra**

* Include a set of *basic* and *derived* set-theoretic operations.
* Procedural: specify a sequence of operations.
	+ performance optimization is important.
* Operations can be unary or binary.
* The result is also a relation: *closure* property => chained operations.
* RA solutions are algorithmic.

***Example:***

See [toyu\_Ex.pdf](https://dcm.uhcl.edu/yue/courses/joinDB/Fall2024/notes/rarc/toyu_Ex.pdf) and [toyu\_RA\_sol\_even.pdf](https://dcm.uhcl.edu/yue/courses/joinDB/Fall2024/notes/rarc/toyu_RA_sol_even.pdf)

**3. Introduction to Relational Calculus**

* Non-procedural, *declarative*, and high level.
* Two kinds:
	1. Domain Relational Calculus (DRC)
	2. Tuple Relational Calculus (TRC)
* Results specified by the *set builder form*: {s | cond(s)}
* cond(s) is known as a *formula*.
* Constructs:
	1. Variables:
		1. TRC: tuples (bound to tuples): e.g,. s, t, student, class, etc.
		2. DRC: Attributes (bound to domain value): e.g., a, b, c, stuId, fname, etc.
		3. RC's variables are sometime known as 'dummy variables'.
	2. Constants: string, int, etc. E.g., 12, 'csci', 3.7.
	3. Comparison operators: <, <, =, etc.
	4. Boolean operators: and (conjunction, ∧ or just ,), or (disjunction ∨), not (¬), implies (⇒), etc.
	5. Membership functions: belongs to, ∈, not belongs to, ∉, etc.
	6. Quantifiers: there exists (existential, ∃), for all (universal ∀).

***Example:***

See [toyu\_Ex.pdf](https://dcm.uhcl.edu/yue/courses/joinDB/Fall2024/notes/rarc/toyu_Ex.pdf) and [toyu\_RC\_sol\_odd.pdf](https://dcm.uhcl.edu/yue/courses/joinDB/Fall2024/notes/rarc/toyu_RC_sol_odd.pdf)

**4. Relational Algebra**

**4.1 Basic Operations**

**4.1.1 Select**

1. Unary operation.
2. Select tuples (with the same schema) based on a Boolean condition.
3. Conditions may include attributes in the relational schema.
4. The Boolean expression of the condition can be composite (containing Boolean expressions joining by logical operators).
5. *'Horizontal' subset*.
6. Not to be confused with the Select statement in SQL.



σcond(R) = {t | t ∈ R and cond}, or simply

σcond(R) = {t | t ∈ R, cond}

 2. Show all information of students majoring in 'CSCI'.

**Example:** All information of students majoring in CSCI.

σmajor='CSCI'(Student)

+--------+-------+-------+-------+-------+---------+---------+
| stuId  | fname | lname | major | minor | credits | advisor |
+--------+-------+-------+-------+-------+---------+---------+
| 100000 | Tony  | Hawk  | CSCI  | CINF  |      40 |    1011 |
| 100001 | Mary  | Hawk  | CSCI  | CINF  |      35 |    1011 |
| 100002 | David | Hawk  | CSCI  | ITEC  |      66 |    1011 |
+--------+-------+-------+-------+-------+---------+---------+
3 rows

In SQL, this is just:

SELECT \*
FROM Student
WHERE major = 'CSCI';

**4.1.2 Project**

1. Unary operation
2. Select attributes from tuples.
3. Duplicate results removed (because a relation is a set).
4. 'Vertical' subset.

πc1, .., cm(R) = {s | ∃t ∈ R (t(ci) = s(ci), for 1 <= i <= m)},

or simply

πc1, .., cm(R) = {s | t ∈ R (t(ci) = s(ci), for 1 <= i <= m)}



**Example:** Names and majors of students

πLName, FName, Major(Student):

SQL:

SELECT LName, FName, Major

FROM student;

+-----------+---------+-------+
| FName     | LName   | Major |
+-----------+---------+-------+
| Tony      | Hawk    | CSCI  |
| Mary      | Hawk    | CSCI  |
| David     | Hawk    | CSCI  |
| Catherine | Lim     | ITEC  |
| Larry     | Johnson | ITEC  |
| Linda     | Johnson | CINF  |
| Lillian   | Johnson | CINF  |
| Ben       | Zico    | NULL  |
| Bill      | Ching   | ARTS  |
| Linda     | King    | ARTS  |
+-----------+---------+-------+
10 rows

 4. Show the names and credits of students majoring in 'CSCI' and having 40 or more credits.

(4) π fName, lName, credit(σMajor=’CSCI’ and credit >= 40(student))

**4.1.3.Cartesian Product**

1. Same as the usual definition of the Cartesian Product of two sets.
	1. Remember that a relation is a set.
2. Merge all possible information from two relations.
3. Also called Cross Product or Cross Join.
4. Name ambiguity may be resolved by using full names.
5. The cardinality of a set S is |S|, the number of elements in the set.
6. |RxS|= |R| \* |S|
7. Not very useful in practice as the result can be large and constructing the result can be time consuming.



{1,2} \* {x,y} = { (1,x), (1,y), (2,x), (2,y)}

***Example:***

R(A,B,C) has three tuples. S(A,D) has four tuples.

The result of R \* S always has 12 tuples with the schema (R.A, B, C, S.A, D).

**Example:** in toyu

**SELECT** s.\*, e.\*

**FROM** student **AS** S **CROSS** **JOIN** enroll **AS** e;

* + - Student \* enroll;theta

student:
+--------+-----------+---------+-------+-------+---------+---------+
| stuId  | fname     | lname   | major | minor | credits | advisor |
+--------+-----------+---------+-------+-------+---------+---------+
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 |
+--------+-----------+---------+-------+-------+---------+---------+
10 rows in set

enroll:
+--------+---------+-------+----------+
| stuId  | classId | grade | n\_alerts |
+--------+---------+-------+----------+
| 100000 |   10000 | A     |        0 |
| 100001 |   10000 | NULL  |     NULL |
| 100002 |   10000 | B-    |        3 |
| 100000 |   10001 | A     |        2 |
| 100001 |   10001 | A-    |        0 |
| 100000 |   10002 | B+    |        1 |
| 100002 |   10002 | B+    |        2 |
| 100000 |   10003 | C     |        0 |
| 100002 |   10003 | D     |        4 |
| 100004 |   10003 | A     |        0 |
| 100005 |   10003 | NULL  |     NULL |
| 100000 |   10004 | A-    |        1 |
| 100004 |   10004 | B+    |     NULL |
| 100005 |   10004 | A-    |        0 |
| 100006 |   10004 | C+    |     NULL |
| 100005 |   10005 | A-    |        0 |
| 100006 |   10005 | A     |     NULL |
| 100005 |   10006 | B+    |     NULL |
| 100007 |   10007 | F     |        4 |
| 100008 |   10007 | C-    |        0 |
| 100007 |   10008 | A-    |        0 |
| 100000 |   11001 | D     |        4 |
+--------+---------+-------+----------+
22 rows

student \* enroll:
+--------+-----------+---------+-------+-------+---------+---------+--------+---------+-------+----------+
| stuId  | fname     | lname   | major | minor | credits | advisor | stuId  | classId | grade | n\_alerts |
+--------+-----------+---------+-------+-------+---------+---------+--------+---------+-------+----------+
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10000 | A     |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100000 |   10000 | A     |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100000 |   10000 | A     |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100000 |   10000 | A     |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100000 |   10000 | A     |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100000 |   10000 | A     |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100000 |   10000 | A     |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100000 |   10000 | A     |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100000 |   10000 | A     |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100000 |   10000 | A     |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100001 |   10000 | NULL  |     NULL |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100001 |   10000 | NULL  |     NULL |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100001 |   10000 | NULL  |     NULL |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100001 |   10000 | NULL  |     NULL |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100001 |   10000 | NULL  |     NULL |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100001 |   10000 | NULL  |     NULL |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100001 |   10000 | NULL  |     NULL |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100001 |   10000 | NULL  |     NULL |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100001 |   10000 | NULL  |     NULL |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100001 |   10000 | NULL  |     NULL |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100002 |   10000 | B-    |        3 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100002 |   10000 | B-    |        3 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100002 |   10000 | B-    |        3 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100002 |   10000 | B-    |        3 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100002 |   10000 | B-    |        3 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100002 |   10000 | B-    |        3 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100002 |   10000 | B-    |        3 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100002 |   10000 | B-    |        3 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100002 |   10000 | B-    |        3 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100002 |   10000 | B-    |        3 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10001 | A     |        2 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100000 |   10001 | A     |        2 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100000 |   10001 | A     |        2 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100000 |   10001 | A     |        2 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100000 |   10001 | A     |        2 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100000 |   10001 | A     |        2 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100000 |   10001 | A     |        2 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100000 |   10001 | A     |        2 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100000 |   10001 | A     |        2 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100000 |   10001 | A     |        2 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100001 |   10001 | A-    |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100001 |   10001 | A-    |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100001 |   10001 | A-    |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100001 |   10001 | A-    |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100001 |   10001 | A-    |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100001 |   10001 | A-    |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100001 |   10001 | A-    |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100001 |   10001 | A-    |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100001 |   10001 | A-    |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100001 |   10001 | A-    |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10002 | B+    |        1 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100000 |   10002 | B+    |        1 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100000 |   10002 | B+    |        1 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100000 |   10002 | B+    |        1 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100000 |   10002 | B+    |        1 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100000 |   10002 | B+    |        1 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100000 |   10002 | B+    |        1 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100000 |   10002 | B+    |        1 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100000 |   10002 | B+    |        1 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100000 |   10002 | B+    |        1 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100002 |   10002 | B+    |        2 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100002 |   10002 | B+    |        2 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100002 |   10002 | B+    |        2 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100002 |   10002 | B+    |        2 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100002 |   10002 | B+    |        2 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100002 |   10002 | B+    |        2 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100002 |   10002 | B+    |        2 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100002 |   10002 | B+    |        2 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100002 |   10002 | B+    |        2 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100002 |   10002 | B+    |        2 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10003 | C     |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100000 |   10003 | C     |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100000 |   10003 | C     |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100000 |   10003 | C     |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100000 |   10003 | C     |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100000 |   10003 | C     |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100000 |   10003 | C     |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100000 |   10003 | C     |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100000 |   10003 | C     |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100000 |   10003 | C     |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100002 |   10003 | D     |        4 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100002 |   10003 | D     |        4 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100002 |   10003 | D     |        4 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100002 |   10003 | D     |        4 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100002 |   10003 | D     |        4 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100002 |   10003 | D     |        4 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100002 |   10003 | D     |        4 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100002 |   10003 | D     |        4 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100002 |   10003 | D     |        4 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100002 |   10003 | D     |        4 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100004 |   10003 | A     |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100004 |   10003 | A     |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100004 |   10003 | A     |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100004 |   10003 | A     |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100004 |   10003 | A     |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100004 |   10003 | A     |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100004 |   10003 | A     |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100004 |   10003 | A     |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100004 |   10003 | A     |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100004 |   10003 | A     |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100005 |   10003 | NULL  |     NULL |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100005 |   10003 | NULL  |     NULL |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100005 |   10003 | NULL  |     NULL |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100005 |   10003 | NULL  |     NULL |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100005 |   10003 | NULL  |     NULL |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10003 | NULL  |     NULL |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100005 |   10003 | NULL  |     NULL |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100005 |   10003 | NULL  |     NULL |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100005 |   10003 | NULL  |     NULL |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100005 |   10003 | NULL  |     NULL |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10004 | A-    |        1 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100000 |   10004 | A-    |        1 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100000 |   10004 | A-    |        1 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100000 |   10004 | A-    |        1 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100000 |   10004 | A-    |        1 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100000 |   10004 | A-    |        1 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100000 |   10004 | A-    |        1 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100000 |   10004 | A-    |        1 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100000 |   10004 | A-    |        1 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100000 |   10004 | A-    |        1 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100004 |   10004 | B+    |     NULL |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100004 |   10004 | B+    |     NULL |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100004 |   10004 | B+    |     NULL |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100004 |   10004 | B+    |     NULL |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100004 |   10004 | B+    |     NULL |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100004 |   10004 | B+    |     NULL |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100004 |   10004 | B+    |     NULL |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100004 |   10004 | B+    |     NULL |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100004 |   10004 | B+    |     NULL |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100004 |   10004 | B+    |     NULL |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100005 |   10004 | A-    |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100005 |   10004 | A-    |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100005 |   10004 | A-    |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100005 |   10004 | A-    |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100005 |   10004 | A-    |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10004 | A-    |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100005 |   10004 | A-    |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100005 |   10004 | A-    |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100005 |   10004 | A-    |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100005 |   10004 | A-    |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100006 |   10004 | C+    |     NULL |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100006 |   10004 | C+    |     NULL |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100006 |   10004 | C+    |     NULL |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100006 |   10004 | C+    |     NULL |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100006 |   10004 | C+    |     NULL |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100006 |   10004 | C+    |     NULL |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100006 |   10004 | C+    |     NULL |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100006 |   10004 | C+    |     NULL |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100006 |   10004 | C+    |     NULL |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100006 |   10004 | C+    |     NULL |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100005 |   10005 | A-    |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100005 |   10005 | A-    |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100005 |   10005 | A-    |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100005 |   10005 | A-    |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100005 |   10005 | A-    |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10005 | A-    |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100005 |   10005 | A-    |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100005 |   10005 | A-    |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100005 |   10005 | A-    |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100005 |   10005 | A-    |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100006 |   10005 | A     |     NULL |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100006 |   10005 | A     |     NULL |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100006 |   10005 | A     |     NULL |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100006 |   10005 | A     |     NULL |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100006 |   10005 | A     |     NULL |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100006 |   10005 | A     |     NULL |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100006 |   10005 | A     |     NULL |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100006 |   10005 | A     |     NULL |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100006 |   10005 | A     |     NULL |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100006 |   10005 | A     |     NULL |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100005 |   10006 | B+    |     NULL |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100005 |   10006 | B+    |     NULL |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100005 |   10006 | B+    |     NULL |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100005 |   10006 | B+    |     NULL |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100005 |   10006 | B+    |     NULL |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10006 | B+    |     NULL |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100005 |   10006 | B+    |     NULL |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100005 |   10006 | B+    |     NULL |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100005 |   10006 | B+    |     NULL |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100005 |   10006 | B+    |     NULL |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100007 |   10007 | F     |        4 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100007 |   10007 | F     |        4 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100007 |   10007 | F     |        4 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100007 |   10007 | F     |        4 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100007 |   10007 | F     |        4 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100007 |   10007 | F     |        4 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100007 |   10007 | F     |        4 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100007 |   10007 | F     |        4 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100007 |   10007 | F     |        4 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100007 |   10007 | F     |        4 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100008 |   10007 | C-    |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100008 |   10007 | C-    |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100008 |   10007 | C-    |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100008 |   10007 | C-    |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100008 |   10007 | C-    |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100008 |   10007 | C-    |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100008 |   10007 | C-    |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100008 |   10007 | C-    |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100008 |   10007 | C-    |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100008 |   10007 | C-    |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100007 |   10008 | A-    |        0 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100007 |   10008 | A-    |        0 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100007 |   10008 | A-    |        0 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100007 |   10008 | A-    |        0 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100007 |   10008 | A-    |        0 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100007 |   10008 | A-    |        0 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100007 |   10008 | A-    |        0 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100007 |   10008 | A-    |        0 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100007 |   10008 | A-    |        0 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100007 |   10008 | A-    |        0 |
| 100000 | Tony      | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   11001 | D     |        4 |
| 100001 | Mary      | Hawk    | CSCI  | CINF  |      35 |    1011 | 100000 |   11001 | D     |        4 |
| 100002 | David     | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100000 |   11001 | D     |        4 |
| 100003 | Catherine | Lim     | ITEC  | CINF  |      20 |    1017 | 100000 |   11001 | D     |        4 |
| 100004 | Larry     | Johnson | ITEC  | NULL  |      66 |    1017 | 100000 |   11001 | D     |        4 |
| 100005 | Linda     | Johnson | CINF  | ENGL  |      13 |    1015 | 100000 |   11001 | D     |        4 |
| 100006 | Lillian   | Johnson | CINF  | ITEC  |      18 |    1015 | 100000 |   11001 | D     |        4 |
| 100007 | Ben       | Zico    | NULL  | NULL  |      16 |    NULL | 100000 |   11001 | D     |        4 |
| 100008 | Bill      | Ching   | ARTS  | ENGL  |      90 |    1018 | 100000 |   11001 | D     |        4 |
| 100009 | Linda     | King    | ARTS  | CSCI  |     125 |    1018 | 100000 |   11001 | D     |        4 |
+--------+-----------+---------+-------+-------+---------+---------+--------+---------+-------+----------+
220 row

R \* S in SQL:

SELECT R.\*, S.\*
FROM R, S; -- note that there is no join condition.

**4.1.4 Union**

1. The set union operator.
2. Condition for R U S: R and S must be *union compatible*. Their relation schema must have compatible schema with the same structures. Each corresponding attribute must have the same types (domains).

R U S = {t | t ∈ R V t ∈ S}

**Example:**

Suppose StaffID and FacultyID are union compatible.

 πStaffID(Staff) U πFacultyID(Faculty)

**Example:** All information of students majoring in CSCI or ARTS.

σ(major='CSCI') (Student) U σ(major='ARTS') (Student)

or

σ(major='CSCI') V (major='ARTS') (Student)

+--------+-------+-------+-------+-------+---------+---------+
| stuId  | fname | lname | major | minor | credits | advisor |
+--------+-------+-------+-------+-------+---------+---------+
| 100008 | Bill  | Ching | ARTS  | ENGL  |      90 |    1018 |
| 100009 | Linda | King  | ARTS  | CSCI  |     125 |    1018 |
| 100000 | Tony  | Hawk  | CSCI  | CINF  |      40 |    1011 |
| 100001 | Mary  | Hawk  | CSCI  | CINF  |      35 |    1011 |
| 100002 | David | Hawk  | CSCI  | ITEC  |      66 |    1011 |
+--------+-------+-------+-------+-------+---------+---------+
5 rows

**4.1.5 Difference (Minus)**

1. The set difference operator. {1,2} – {1,3} = {2}
2. R - S: R and S must be *union compatible*.

R - S = {t | t ∈ R and not (t ∈ S)}

or

R - S = {t | t ∈ R, t ∉ S}

**Example:** Information of all students majoring in CSCI but not those taken credits less than 40.

σmajor='CSCI'(Student) - σcredit <40 (Student)

+--------+-------+-------+-------+-------+---------+---------+
| stuId  | fname | lname | major | minor | credits | advisor |
+--------+-------+-------+-------+-------+---------+---------+
| 100000 | Tony  | Hawk  | CSCI  | CINF  |      40 |    1011 |
| 100002 | David | Hawk  | CSCI  | ITEC  |      66 |    1011 |
+--------+-------+-------+-------+-------+---------+---------+
2 rows

Note that this is the same as:

σmajor='CSCI' and credit >=40(Student)

**4.1.6 Rename**

1. Rename the names of selected attributes in a relation.
2. Maybe used to rename attributes before a set operation.
3. Notation in Elmarsi (a popular db textbook):



* A better notation includes the original name and the new name.

Example:

ρ(FacultyId, department <- FacId, deptCode) (Faculty)

+-----------+----------+----------+------------+---------------------+
| facultyId | fname    | lname    | department | rank                |
+-----------+----------+----------+------------+---------------------+
|      1011 | Paul     | Smith    | CSCI       | Professor           |
|      1012 | Mary     | Tran     | CSCI       | Associate Professor |
|      1013 | David    | Love     | CSCI       | NULL                |
|      1014 | Sharon   | Mannes   | CSCI       | Assistant Professor |
|      1015 | Daniel   | Kim      | CINF       | Professor           |
|      1016 | Andrew   | Byre     | CINF       | Associate Professor |
|      1017 | Deborah  | Gump     | ITEC       | Professor           |
|      1018 | Art      | Allister | ARTS       | Assistant Professor |
|      1019 | Benjamin | Yu       | ITEC       | Lecturer            |
|      1020 | Katrina  | Bajaj    | ENGL       | Lecturer            |
|      1021 | Jorginlo | Neymar   | ACCT       | Assistant Professor |
+-----------+----------+----------+------------+---------------------+
11 rows

* The basic set of operations is *complete*. Other relational algebra operations can be derived from them.

12. Show the student names and their major names for all students who have received a grade A in a class offered by a faculty from the CSCI department.

(12) π fName, lName, deptName(student |x| ρmajor <- deptCode(department) |x| σgrade=’A’(enroll) |x| class |x| σdeptCode=’CSCI’(faculty))

ρmajor <- deptCode(department): rename deptCode to major (a common attribute of student)

**4.2. Derived Operations**

**4.2.1 Theta-join**

1. Allow the application of condition on Cartesian product.
2. There are still redundant data on common attributes.
3. Allow the query engine to throw away tuples not in the result immediately.
4. Conceptually, a Cartesian Product followed by a selection Θ.
5. Not usually used.

R1 ⋈ΘR2 = σΘ(R1 \* R2)

**Example:** All related information of students with 70 or more credits and a grade A or better in some courses.

Student ⋈(credits >= 70 and grade = 'A') Enroll

-- theta join

SELECT s.\*, e.\*

FROM student AS S CROSS JOIN enroll AS e ON (ach >= 70 and grade = 'A');

+--------+-------+-------+-------+-------+---------+---------+--------+---------+-------+----------+
| stuId  | fname | lname | major | minor | credits | advisor | stuId  | classId | grade | n\_alerts |
+--------+-------+-------+-------+-------+---------+---------+--------+---------+-------+----------+
| 100008 | Bill  | Ching | ARTS  | ENGL  |      90 |    1018 | 100000 |   10000 | A     |        0 |
| 100008 | Bill  | Ching | ARTS  | ENGL  |      90 |    1018 | 100000 |   10001 | A     |        2 |
| 100008 | Bill  | Ching | ARTS  | ENGL  |      90 |    1018 | 100004 |   10003 | A     |        0 |
| 100008 | Bill  | Ching | ARTS  | ENGL  |      90 |    1018 | 100006 |   10005 | A     |     NULL |
| 100009 | Linda | King  | ARTS  | CSCI  |     125 |    1018 | 100000 |   10000 | A     |        0 |
| 100009 | Linda | King  | ARTS  | CSCI  |     125 |    1018 | 100000 |   10001 | A     |        2 |
| 100009 | Linda | King  | ARTS  | CSCI  |     125 |    1018 | 100004 |   10003 | A     |        0 |
| 100009 | Linda | King  | ARTS  | CSCI  |     125 |    1018 | 100006 |   10005 | A     |     NULL |
+--------+-------+-------+-------+-------+---------+---------+--------+---------+-------+----------+
8 rows in set (0.01 sec)

 **4.2.2 Equi-join**

1. Theta-join where the condition involves only equality comparisons.
2. There are still redundant data on *common attributes*.
3. Common attributes are attributes that have the same *names*. The attributes may not have the same meaning.
4. Not usually used.

**Example:**

Student |x| (Student.StuId = Enrol.StuId) Enroll

+--------+---------+---------+-------+-------+---------+---------+--------+---------+-------+----------+
| stuId  | fname   | lname   | major | minor | credits | advisor | stuId  | classId | grade | n\_alerts |
+--------+---------+---------+-------+-------+---------+---------+--------+---------+-------+----------+
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10000 | A     |        0 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10001 | A     |        2 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10002 | B+    |        1 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10003 | C     |        0 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   10004 | A-    |        1 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 | 100000 |   11001 | D     |        4 |
| 100001 | Mary    | Hawk    | CSCI  | CINF  |      35 |    1011 | 100001 |   10000 | NULL  |     NULL |
| 100001 | Mary    | Hawk    | CSCI  | CINF  |      35 |    1011 | 100001 |   10001 | A-    |        0 |
| 100002 | David   | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100002 |   10000 | B-    |        3 |
| 100002 | David   | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100002 |   10002 | B+    |        2 |
| 100002 | David   | Hawk    | CSCI  | ITEC  |      66 |    1011 | 100002 |   10003 | D     |        4 |
| 100004 | Larry   | Johnson | ITEC  | NULL  |      66 |    1017 | 100004 |   10003 | A     |        0 |
| 100004 | Larry   | Johnson | ITEC  | NULL  |      66 |    1017 | 100004 |   10004 | B+    |     NULL |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10003 | NULL  |     NULL |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10004 | A-    |        0 |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10005 | A-    |        0 |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 | 100005 |   10006 | B+    |     NULL |
| 100006 | Lillian | Johnson | CINF  | ITEC  |      18 |    1015 | 100006 |   10004 | C+    |     NULL |
| 100006 | Lillian | Johnson | CINF  | ITEC  |      18 |    1015 | 100006 |   10005 | A     |     NULL |
| 100007 | Ben     | Zico    | NULL  | NULL  |      16 |    NULL | 100007 |   10007 | F     |        4 |
| 100007 | Ben     | Zico    | NULL  | NULL  |      16 |    NULL | 100007 |   10008 | A-    |        0 |
| 100008 | Bill    | Ching   | ARTS  | ENGL  |      90 |    1018 | 100008 |   10007 | C-    |        0 |
+--------+---------+---------+-------+-------+---------+---------+--------+---------+-------+----------+
22 rows

It is important to note the difference between names and meanings. Consider

student(stuId, ... advisorFacId, ..., createTime) and

faculty(facId, ..., createTime)

1. The attributes createTime in student and faculty have the same name, but different meaning.
	1. Student(createTime) is the time the student row is inserted into the student table.
	2. Faculty(createTime) is the time the faculty row is inserted into the faculty table
2. student(advisorFacId) and faculty(facId) have different names but the same meaning. In fact, student(advisorFacId) is a foreign key that references faculty(facId).

**4.2.3 Natural Join**

1. Remove redundant common attributes from equi-joins
	1. Equi-join on all common attributes.
	2. Projection to remove redundant common attributes.
2. Used very frequently to combine two tables.
3. If two relations do not share any common attributes, their natural join is the same as their Cartesian Product.

Let C1, C2, ... Cm be the common attributes of R and S.

R |x| S = πA1, A2, .. Al(σR.C1=S.C1,.., R.Cm=S.Cm(R\*S))

where A1, A2, ... Al is the list of attributes in R\*S except S.C1, S.C2,.. S.Cm.

**Example:**

The schema of R(A,B) |x| S(A,C) is ABC. The schema of R(A,B) \* S(A,C) is {R.A, B, S.A, C}.

1. Show the names of students who have enrolled in 10000.

(8)

π fName, lName(student |x| σclassId = 10000(enroll)) -- [1]

π fName, lName(student |x| σclassId = 10000(enroll)) -- [1] 3 operations; size of the intermediate results (3 rows): small

π fName, lName(σclassId = 10000(student |x| enroll)) -- [2] 3 operations; size of the intermediate results (242 rows): large

π fName, lName(student |x| π STuId (σclassId = 10000(enroll))) -- [3] 4 operations; size of the intermediate results:

smallest

Query optimization: minimize (1) # of operations; (2) size of the intermediate results:

**Example:**

Student |x| Enroll:

+--------+---------+---------+-------+-------+---------+---------+---------+-------+----------+
| stuId  | fname   | lname   | major | minor | credits | advisor | classId | grade | n\_alerts |
+--------+---------+---------+-------+-------+---------+---------+---------+-------+----------+
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 |   10000 | A     |        0 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 |   10001 | A     |        2 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 |   10002 | B+    |        1 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 |   10003 | C     |        0 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 |   10004 | A-    |        1 |
| 100000 | Tony    | Hawk    | CSCI  | CINF  |      40 |    1011 |   11001 | D     |        4 |
| 100001 | Mary    | Hawk    | CSCI  | CINF  |      35 |    1011 |   10000 | NULL  |     NULL |
| 100001 | Mary    | Hawk    | CSCI  | CINF  |      35 |    1011 |   10001 | A-    |        0 |
| 100002 | David   | Hawk    | CSCI  | ITEC  |      66 |    1011 |   10000 | B-    |        3 |
| 100002 | David   | Hawk    | CSCI  | ITEC  |      66 |    1011 |   10002 | B+    |        2 |
| 100002 | David   | Hawk    | CSCI  | ITEC  |      66 |    1011 |   10003 | D     |        4 |
| 100004 | Larry   | Johnson | ITEC  | NULL  |      66 |    1017 |   10003 | A     |        0 |
| 100004 | Larry   | Johnson | ITEC  | NULL  |      66 |    1017 |   10004 | B+    |     NULL |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 |   10003 | NULL  |     NULL |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 |   10004 | A-    |        0 |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 |   10005 | A-    |        0 |
| 100005 | Linda   | Johnson | CINF  | ENGL  |      13 |    1015 |   10006 | B+    |     NULL |
| 100006 | Lillian | Johnson | CINF  | ITEC  |      18 |    1015 |   10004 | C+    |     NULL |
| 100006 | Lillian | Johnson | CINF  | ITEC  |      18 |    1015 |   10005 | A     |     NULL |
| 100007 | Ben     | Zico    | NULL  | NULL  |      16 |    NULL |   10007 | F     |        4 |
| 100007 | Ben     | Zico    | NULL  | NULL  |      16 |    NULL |   10008 | A-    |        0 |
| 100008 | Bill    | Ching   | ARTS  | ENGL  |      90 |    1018 |   10007 | C-    |        0 |
+--------+---------+---------+-------+-------+---------+---------+---------+-------+----------+
22 rows

**Exercise:**

Let the cardinality of R(A,B) be 5 and the cardinality of S(A,C) be 6. What is the range of the cardinality of R(A,B) |x| S(A,C)?

**4.2.4 Other Joins** (Additional Materials)

1. Some other joins are left join, right join, outer join, inner join and semi-join.
2. They can be defined through relational algebra expressions based on the basic operations.
3. Look them up when needs arise. For example: <https://en.wikipedia.org/wiki/Relational_algebra>

**4.2.5 Division** (Additional Materials)

1. R / S or R ÷ S.
2. Condition: the domain of S is a proper subset of R.
3. Let the schemes of R, S and T be dom(R), dom(S) and dom(T) = dom(R) - dom(S) respectively.
4. R / S = {t | t ∈ dom(T), ꓯs ∈ S (ꓱr ∈ R (r = st))}.
5. In term of basic RA operations, R / S = πR-S(R) - πR-S((πR-S(R) \* S) - R)

**Example:**

Find the student id of all students who enrolled in all courses offered by the faculty '1014':

Stuid and classNumber information (who is enrolled in which class):

π(stuId, classId) (Enroll): rows added to Class.

+--------+---------+
| stuId  | classId |
+--------+---------+
| 100000 |   10000 |
| 100000 |   10001 |
| 100000 |   10002 |
| 100000 |   10003 |
| 100000 |   10004 |
| 100000 |   11001 |
| 100001 |   10000 |
| 100001 |   10001 |
| 100002 |   10000 |
| 100002 |   10002 |
| 100002 |   10003 |
| 100004 |   10003 |
| 100004 |   10004 |
| 100005 |   10003 |
| 100005 |   10004 |
| 100005 |   10005 |
| 100005 |   10006 |
| 100006 |   10004 |
| 100006 |   10005 |
| 100007 |   10007 |
| 100007 |   10008 |
| 100008 |   10007 |
+--------+---------+
22 rows

Classes offered by faculty '1014':

π(classId) (σ(facId='1014) (Class)):

+---------+
| classId |
+---------+
|   10003 |
|   10004 |
+---------+
2 rows

Solution:

π(stuId, classId) (Enroll) / π(stuId, classId) (Enroll):

+--------+
| stuId  |
+--------+
| 100000 |
| 100004 |
| 100005 |
+--------+
3 rows

**4.3 Query Optimization**

* Since RA is operational and thus algorithmic, there are multiple solutions with varying performance.
* Some heuristics for constructing effective RA solutions:
	1. Minimize the number of RA operations.
	2. Minimize the sizes of the intermediate results.
* In SQL:
	1. A SQL query execution plan breaks down a query into basic execution steps (based on RA).
	2. A SQL optimizer selects one from a list of execution plans.
	3. In MySQL, use the EXPLAIN statement to obtain information about the execution plan: <https://dev.mysql.com/doc/refman/8.1/en/explain-output.html>.

**4.4 Epilog**

Some shortcomings of Relational Algebra:

1. Cannot navigate tuples.
2. Cannot deal with recursion.
	1. e.g., for the relation Employee(SSN, Supervisor\_SSN, ...), find all supervisors (direct or indirect).
	2. May extend to logical databases, e.g. Datalog.
3. No group functions.
	1. e.g., Show the available total quantities of all parts.
4. Operations are too simple, resulting in long sequences.

**5. Relational Calculus

5.1 Review**

* Non-procedural, *declarative*, and high level.
* Two kinds:
	1. Domain Relational Calculus (DRC)
	2. Tuple Relational Calculus (TRC)
* Results specified by the *set builder form*: {s | cond(s)}
* cond(s) is known as a *formula*.
* Constructs:
	1. Variables:
		1. TRC: tuples (bound to tuples): e.g,. s, t, student, class, etc.
		2. DRC: Attributes (bound to domain value): e.g., a, b, c, stuId, fname, etc.
		3. RC's variable is sometime known as 'dummy variable'.
	2. Constants: string, int, etc., E.g., 12, 'csci', 3.7.
	3. Comparison operators: <, <, =, etc.
	4. Boolean operators: and (conjunction, ∧ or just ,), or (disjunction ∨), not (¬), implies (⇒), etc.
	5. Membership functions: belongs to, ∈, not belongs to, ∉, etc.
	6. Quantifiers: there exists (existential, ∃), for all (universal ∀).

**5.2 More RC**

* An *atom* can be thought of as a simple Boolean expression:
	+ e ∈ R, or
	+ x op y where x and y are attributes or constants, and op is a comparison operation.
* A *formula* is either an atom or formula connected by Boolean operator or qualifiers.
* A formula that is not an atom can be thought of a compound Boolean expression.
* A variable is *bound* if it appears in qualifier expressions. Otherwise, it is a *free* variable.
* Free variables can only appear in the LHS of |.
* All RA expressions can be expressed in RC.
* RA and RC have the same expressive power.
* Any query language that can express all RA is known to be relational complete.
* Relational Calculus expressions need to be *safe*: results should be a*finite* set of tuples.
* . {s | s ∈ Student }: safe
* Care should be taken especially for the negation operation. E.g. {s |¬ (s ∈ Student) } is unsafe.
* For a given implementation of relational calculus:
	+ There may be restrictions in supported constructs.
	+ There may be certain *canonical (good and supported)* requirements: e.g. *conjunction* (joined by the and operator) of *disjunction* (joined by the or operator).
* 1. Show all student names.
* TRC: variables are tuples /rows; SQL
* {(s.fName, s.lName) | s ϵ student}
* DRC: variables are columns/attributes (values in domains)

{(fName, lName) | (\_, lName, fName, \_, \_, \_, \_) ϵ student}; \_: anonymous variables/placeholders.

{(fName, lName) | (stuId, lName, fName, major, minor, ach, advisor) ϵ student} – 7 variables: columns/attributes

**Example:**

{i | i ∈ I ∧ i % 2 =0}
{i | i ∈ I, i % 2 =0} -- set builder form.

{t | ∃r ∈R, r.firstname = t.firstname, r.lastname = t.lastname}

* t is a *free* variable.
* It will have two attributes: t.firstname and t.lastname.

Alternatively, we can use the set builder form in the LHS before |:

{(r.firstname, r.lastname) | r ∈ R}

R(A,B,C,D) / S(C,D)

{(a,b) | (∀(c,d) ∈ S) (a,b,c,d) ∈ R)}

***Exercises:***

How do you use RC to implement RA operations?

**5.3 TRC**

* The variables in TRC are tuples.
* SQL is based on TRC.

**5.4. DRC**

* The variables in DRC are attributes (domain values).
* Query By Example (QBE) is based on DRC.

**Exercise:**

Work on some of the query questions listed in the [toyu Query Exercise](https://dcm.uhcl.edu/yue/courses/joinDB/Fall2024/notes/rarc/toyu_Ex.pdf) in DRC and TRC.