University of Toronto
csc343, Fall 2019

Assignment 1: Your name and student number here

Delete this before you hand in

Unary operators on relations:

- $\Pi_{x,y,z}(R)$

- $\sigma_{condition}(R)$

- $\rho_{New}(R)$

Binary operators on relations:

- $R \times S$

- $R \bowtie S$

- $R \bowtie_{condition} S$

- $R \cup S$

- $R \cap S$

- $R - S$

Logical operators:

- $\vee$

- $\wedge$

- $\neg$

Assignment:

- $New(a, b, c) := R$

General LaTeX tips:

- The characters "\\" create a line break and "[5pt]" puts in five points of extra vertical space. The algebra is easier to read with extra vertical space.

- Some sequences of letters look terrible in "math mode" (inside dollar signs). For example, `$Offer$` comes out as $Offer$. You can use `\var` to fix that. For example, `$\var{Offer}$` comes out as *Offer*.

- When an expression is deeply nested, it can be hard to match the brackets with your eyes. You can vary the size of the brackets to make that easier. This LaTeX:

    `\Bigg(  \bigg[   \Big(  \big(   \big)  \Big)  \bigg]  \Bigg)`

    produces this result:

    $$\Bigg(\bigg[\Big(\big(\big) \Big) \bigg] \Bigg)$$

- Some of your expressions will become very wide because there are multiple, long conditions on a select. You can stack those conditions using \substack. Here's an example:

$$\sigma_{\substack{x<20 \\ \wedge \\ y=10}}(Some\ big\ expression)$$

  Notice that this requires using the package called "amsmath", which is accomplished by saying \usepackage{amsmath}. In the present file, we brought in a couple of other packages and made a couple of definitions (see the top of this file).

Below is the text of the assignment questions; I suggest you include it in your solution. I also included a nonsense example of how a query might look in LaTeX. As required in this assignment, it uses "–" to indicate comments. Note that I have added less vertical space between comments and the algebra they pertain to than between steps in the algebra. This helps the comments visually stick to the algebra.

Part 1: Queries

1. Context: New staff need oversight, particularly volunteers, because learning the Chenhall system takes practise.

   Query: Find the latest object catalogued by the newest volunteer. Report the volunteer's ID number and first name, as well as the object's description, type, and date of cataloguing.

   The possible ties are interesting in this query. For example, two volunteers could be tied for newest. They *each* should be included in the result with their own latest-catalogued object(s), even if one volunteer's latest-catalogued object(s) were catalogued before the other volunteer's latest-catalogued object(s).

   – sID has a grade of at least 85.
   $$HaveHighGrade(sID) := \Pi_{sID}\sigma_{grade \geq 85}Took$$

   – sID passed a course taught by Atwood.
   $$PassedAtwood(sID) := \Pi_{sID}\sigma_{instructor:="Atwood" \wedge grade \geq 50}(Took \bowtie Offering)$$

   – sID got 100 at least twice.
   $$AtLeastTwice(sID) :=$$
   $$\Pi_{T1.sID}\sigma_{T1.oID \neq T2.oID \wedge T1.sID=T2.sID \wedge T1.grade=100 \wedge T2.grade=100}[(\rho_{T1}Took) \times (\rho_{T2}Took)]$$

2. Context: We are curious about donors who've made very broad contributions.

   Query: For each donor who has made a donation in every Chenhall category in the database, report in a single tuple:
   - the value of the single most valuable object they have ever donated, and
   - the value of the single least valuable object they have ever donated.

3. Context: There is no good reason for this query. :-)

   Query: For each donation with three or more objects in it, report the second tallest object in the donation. Report the object's catalogue number, height, and width, as well as the ID of the donation it came from.

4. Context: We are looking for donations that may contain unusual objects.

   Query: Find any donation that was catalogued entirely by one staff person, where the same tag was used for at least 2 objects, but that tag was never used for any object catalogued by the same staff person in any other donation. For each such donation, report the donation ID and donor ID, and the staff member ID of the one person who catalogued it.

5. Context: We want to bring back volunteers who used to help consistently but no longer do.

   Query: Find all volunteers who have catalogued object(s) from at least two different donations every year up to and including in 2016, but have catalogued nothing since. By "every year" we mean every year that appears in the *Object* relation. For each of these volunteers, report their staff ID and email address.

6. Context: For efficiency, it makes sense that a donation with very similar objects be catalogued by a single person.

   Query: A "consistent donation" is one in which every object is from the same Chenhall category. Find all donations that were consistent, yet more that one person catalogued object(s) from it. For each qualifying donation, include in the result one tuple for each person who catalogued objects from it. Each tuple should contain the donation ID and person's staff member ID.

7. Context: Cataloguing a very large donation takes a great deal of diligence. We want to identify and reward staff who have done so.

   Query: The "largest donation" is the one with the most objects in it (there could be ties for largest). Find any donor who made a donation that qualifies as largest. Report the ID of each staff member who has catalogued an object that was part of any donation by any of these donors.

8. Context: We want to identify staff who are not using tags enough, so that we can provide additional training.

   Query: Find staff who have, two or more times, catalogued every object in a donation but gave none of them tags. Report only the SID of these staff members.

9. Context: We don't want staff isolating themselves.

   Query: Find all pairs of staff members who have catalogued together yet neither of them has catalogued with anyone else. (Two people have "catalogued together" if they have each catalogued object(s) from the same donation.) Put each pair into a tuple that includes their staff IDs and their email addresses. Don't include pseudo-duplicates, that is, don't report A, B and also B, A.

10. Context: We are exploring the categorization system.

    Query: A Chenhall category is "complete" if it has at least one primary term and each of its primary terms has at least one secondary term. Find all Chenhall categories that are not complete.

Part 2: Additional Integrity Constraints

1. No secondary term can be a primary term or a Chenhall category, and no primary term can be a Chenhall category.

2. A donation can be catalogued only if (a) it has one object, which is worth at least $100, or (b) it has two or more objects, which are worth at least $150 in total.

3. Each object catalogued before 2018 must have exactly three tags.

4. There are strict rules for who is allowed to catalogue what:

Staff type	Chenhall category		
	"personal artifacts"	"architectural"	any other category
temp	✗	✗	✗
volunteer	✓	✓	✗
intern	✓	✓	✗
permanent	✓	✓	✓