Assignment: Dependency Theory

COMPSCI 2DB3: Databases–Winter 2024

Deadline: March 16, 2024

Department of Computing and Software
McMaster University

Please read the *Course Outline* for the general policies related to assignments.

Plagiarism is a *serious academic offense* and will be handled accordingly.
All suspicions will be reported to the *Office of Academic Integrity*
(in accordance with the Academic Integrity Policy).

This assignment is an *individual* assignment: do not submit work of others. All parts of your submission *must* be your own work and be based on your own ideas and conclusions. Only *discuss or share* any parts of your submissions with your TA or instructor. You are *responsible for protecting* your work: you are strongly advised to password-protect and lock your electronic devices (e.g., laptop) and to not share your logins with partners or friends!

If you *submit* work, then you are certifying that you are aware of the *Plagiarism and Academic Dishonesty* policy of this course outlined in this section, that you are aware of the Academic Integrity Policy, and that you have completed the submitted work entirely yourself. Furthermore, by submitting work, you agree to automated and manual plagiarism checking of all submitted work.

Late submission policy. Late submissions will receive a late penalty of 20% on the score per day late (with a five hour grace period on the first day, e.g., to deal with technical issues) and submissions five days (or more) past the due date are not accepted. In case of technical issues while submitting, contact the instructor *before* the deadline.

1 Functional Dependencies

The Armstrong Axioms are sound, complete, and independent. They are not the only options for reasoning about functional dependencies, however. For example, we have seen the Union and Decomposition rules. Next, we will explorer some alternative options:

Problem 1.

Consider the following inference rules:

1. If $X \longrightarrow Y$ and $V \subseteq (W \cup Y)$, then $XW \longrightarrow V$.
2. If $X \longrightarrow Y$ and $Y \longrightarrow Z$, then $X \longrightarrow XYZ$.
3. $\emptyset \longrightarrow \emptyset$.

P1.1. Prove that the *first* inference rule is sound using only the Armstrong Axioms.

P1.2. Prove that the *second* inference rule is sound directly from the definition of functional dependencies (without using any inference rules).

P1.3. Prove that the *third* inference rule is sound.

P1.4. Are these inference rules complete? If so, then prove why. Otherwise, show why not.

Hint: You may assume that the Armstrong Axioms are complete.

P1.5. Are these inference rules independent? If so, then show why. Otherwise, show why not.

Problem 2.

Consider the following two inference rules:

1. If $X \longrightarrow Y$ and $V \longrightarrow W$, then $XV \longrightarrow YW$.
2. If $X \longrightarrow Y$ and $V \longrightarrow W$ with $V \subseteq Y$, then $X \longrightarrow W$.

P2.1. Prove that these inference rules are sound.

P2.2. Are these inference rules complete? If so, then prove why. Otherwise, show why not.

Hint: You may assume that the Armstrong Axioms are complete.

P2.3. Are these inference rules independent? If so, then show why. Otherwise, show why not.

2 Closure and minimal cover

Problem 3. Consider the relational schema $\mathbf{r}(A, B, C, D, E)$ and the following set of functional dependencies:

$$\mathfrak{S} = \{A \longrightarrow B, BC \longrightarrow D, CD \longrightarrow E, AC \longrightarrow DE, ABD \longrightarrow C\}.$$

P3.1. Provide the attribute closure of set of attributes A (hence, A^+) and of set of attributes AC (hence, $(AC)^+$) with respect to $\mathfrak{S}$. Explain your steps.

P3.2. Compute the closure $\mathfrak{S}^+$. Explain your steps. Based on the closure, indicate which attributes are *superkeys* and which attributes are *keys*.

Hint: A *superkey* is a set of attributes that determines *all* attributes from $\mathbf{r}$. A *key* k is a superkey of minimal size (if we remove any attribute from k, it is no longer a key).

P3.3. Provide a minimal cover for $\mathfrak{S}$. Explain your steps.

3 Multi-valued and inclusion dependencies

Problem 4. Consider the following dependencies for relational schemas R and S:

1. $X \twoheadrightarrow Z \setminus X$ with Z all attributes of the relational schema.
2. If $Y \cap Z = \emptyset$, $X \twoheadrightarrow Y$, and $Z \twoheadrightarrow W$, then $X \twoheadrightarrow W \setminus Y$.
3. If $R[AB] \subseteq S[CD]$ and $C \longrightarrow D$, then $A \longrightarrow B$.
4. $R[X] \subseteq R[X]$.

In the above, $S[CC]$ refers to the relational schema S in which the column C is used twice.

P4.1. For each inference rule: Is the inference rule sound? If so, then prove that the rule is sound. Otherwise, provide a counterexample.

Assignment

The goal of the assignment is to practice dependency theory. To do so, you have to write a report in which you solve each of the above problems. Your submission:

1. must include your student number and MacID;
2. must be a PDF file;
3. must have clearly labeled solutions to each of the stated problems;
4. must include a correct proof for all proofs in problems 1–4;
5. must include a clear counterexample for all counterexamples in problems 1–4;
6. must be clearly presented;
7. must *not* be hand-written: prepare your document in Microsoft Word or another word processor (printed or exported to PDF) or in LaTeX.

Submissions that do not follow the above requirements will get a grade of zero.

Grading

The presented solutions for the problems in Section 1 will account for 50% of the maximum grade; the presented solutions for the problems in Section 2 will account for 30% of the maximum grade; and the presented solution for Section 3 will account for 20% of the maximum grade.