Problem 1 (5 points) Let R be a relation with attributes ABCD. Consider the query

\[ \pi_{AB}\left[\pi_{BCD}(R) \bowtie \pi_{ACD}(R)\right] \bowtie \pi_{AD}(R) \]

(i) (1 point) Construct the tableau corresponding to the query.

(ii) (3 points) Minimize the tableau in (i) knowing that the query is only applied to databases satisfying the dependencies:

\[ B \rightarrow D, D \rightarrow C, A \rightarrow B \]

(iii) (1 point) Construct an algebra query corresponding to the minimized tableau obtained in (ii).

Problem 2 (4 points) Show that there is no algorithm that, given a relational algebra query, returns an equivalent query with a minimum number of algebra operators.

Problem 3 (12 points) A database holds data about courses in an undergraduate CSE program, prerequisites, and records for each student. The tables are the following:

<table>
<thead>
<tr>
<th>Course</th>
<th>C-id</th>
<th>Record</th>
<th>S-id</th>
<th>C-id</th>
<th>Prereq</th>
<th>C-id</th>
<th>P-id</th>
</tr>
</thead>
</table>

In relation Prereq, P-id provides a prerequisite for C-id (assume this is acyclic). Relation Record provides, for each student S-id, the courses C-id he/she has already taken. Assume that Record[C-id] ⊆ Course[C-id], Prereq[C-id] ⊆ Course[C-id] and Prereq[P-id] ⊆ Course[C-id]. Consider the following query:

“For each student, find the set of courses whose prerequisites he/she has already taken”

Express this query in the following languages: (i) CALC (ii) relational algebra (iii) nr-Datalog° (iv) inflationary Datalog°, and (v) semi-positive Datalog°,
assuming also given a successor relation $\text{succ}$ on course ids, as well as constants $\text{min}$ and $\text{max}$ providing the minimum, resp. maximum course id.

**Problem 4** (5 points) Consider a database for metro and bus lines, consisting of two relations $\text{Metro}[	ext{Station}, \text{Next-Station}]$ and $\text{Bus}[	ext{Station}, \text{Next-Station}]$. Write a stratifiable Datalog program to find the pairs of stations $(a, b)$ such that $b$ can be reached from $a$ by some combination of metro and bus, but not by metro or bus alone.

**Problem 5** (2 points) Let $G$ be a binary relation representing the edges of a directed graph. Is there a CALC sentence stating that $G$ consists of a single cycle?

**Problem 6** (3 points) We know that Datalog queries are monotonic and can be evaluated in PTIME (with respect to the database). Does Datalog express all the monotonic PTIME queries?

**Problem 7** (4 points) Prove that it is undecidable whether a \textit{while} $^+$ $\text{W}$ program defines a deterministic query.