1. Explain how a database is an ADT as follows.

   (a) (3 points) What is the set of values? (Be sure to explain what the full set of values is with respect to the DDL, not just the current state.)

   (b) (3 points) What is the set of operations? (Don’t get carried away here, but be sure to give examples of the DML for both updating and retrieving information.)

   (c) (4 points) Explain how the principle of information hiding for ADTs applies.
2. (15 points) Create an ER diagram for the following information.

A PUBLISHER publishes many BOOKs, but a BOOK is published by only one PUBLISHER. A PUBLISHER is uniquely identified by a Name and has offices in several CITYs. A CityName and State together uniquely identifies a CITY. A BOOK has a Title, but the title does not uniquely identify the book. A BOOK, however, does have a unique identifying Number. A BOOK may be adopted by several UNIVERSITYs, each of which is uniquely identified by its UniversityName. A UNIVERSITY also has an Address. We are also interested in the date the BOOK was first adopted by a UNIVERSITY.

Use the notation in our text. Each noun in all capital letters identifies an entity set that must appear in your diagram. No other entity set may appear in your diagram. Each noun with only an initial capital letter identifies an attribute that must appear in your diagram. No other attributes may appear in your diagram. Use the ER notation properly to make relationship sets one-one, one-many, or many-many. Also, use the ER notation to identify keys for entity sets.

3. (12 points) Construct appropriate schemes (table headers) for your ER diagram in problem 2. Underline all minimal keys for each table, and be sure to make a continuous underline for composite keys. Use double underlining for primary keys. Also, for your schemes, explicitly specify all foreign keys by drawing an arrow from a foreign key to the primary key it references.
4. Let \( r \) and \( s \) be relations as follows:

\[
\begin{array}{ccc}
A & B & C \\
\hline
8 & 7 & 5 \\
1 & 3 & 6 \\
1 & 2 & 3 \\
4 & 7 & 3 \\
8 & 9 & 2 \\
\end{array}
\quad s =
\begin{array}{ccc}
B & C & D \\
\hline
6 & 5 & 8 \\
7 & 6 & 5 \\
2 & 3 & 4 \\
9 & 6 & 7 \\
2 & 3 & 5 \\
\end{array}
\]

Compute the relation for each of the following:

(a) (4 points) \( \pi_C s \)

(b) (4 points) \( \sigma_{C=0} s \)

(c) (4 points) \( \pi_B s - \pi_B r \)

(d) (4 points) \( r \bowtie s \)

(e) (4 points) \( (\pi_A r) \bowtie (\pi_{EρC→E} s) \)
5. Consider the following database scheme.

\[
\begin{align*}
\text{yuppie}(SSN, \text{Name}, \text{City}, \text{State}, \text{Salary}) & \quad \text{key}: SSN \\
\text{car}(\text{VIN}, \text{Make}, \text{Model}, \text{Year}, \text{Cost}, \text{Color}) & \quad \text{key}: \text{VIN} \\
\text{owns}(SSN, VIN) & \quad \text{key}: \{SSN, VIN\}
\end{align*}
\]

(a) (6 points) Write a relational-algebra query to find the names of yuppies who own red cars and whose salary is less than $200,000.

(b) (6 points) Write a relational algebra expression to find the cars that cost more than $50,000 but are not red.

(c) (6 points) Convert the following SQL query to relational algebra.

\[
\begin{align*}
\text{select } y1.\text{SSN}, y1.\text{Name}, y2.\text{SSN}, y2.\text{Name} \\
\text{from yuppie y1, yuppie y2} \\
\text{where } y1.\text{Name} = y2.\text{Name} \text{ and } y1.\text{SSN} < y2.\text{SSN}
\end{align*}
\]
6. Let \( r_1(ABC), r_2(CDE), \) and \( r_3(ACF) \), and assume that the domain for all attributes is the set of integers. Optimize the following queries.

(a) (5 points) \( \sigma_{A=4 \land B=6 \land F \geq 7}(r_1 \bowtie r_2 \bowtie r_3) \)

(b) (5 points) \( \pi_A \sigma_{C=5}(\pi_{AC}r_1 \cup \pi_{AC}r_3) \)

(c) (5 points) \( \pi_{EF}(r_2 \bowtie r_3) \)

7. (10 points) Consider the following transactions:

\[
\begin{align*}
T_0: & \quad \text{Read}(A) \quad \text{Read}(A) \\
& \quad A := A + 2 \quad A := A - 1 \\
& \quad \text{Write}(A) \quad \text{Write}(A) \\
& \quad \text{Read}(B) \quad \text{Read}(B) \\
& \quad B := B + 1 \\
& \quad \text{Write}(B)
\end{align*}
\]

Let \( A \) initially be 2, and let \( B \) initially be 5. Assume that \( T_0 \) executes then \( T_1 \) executes. If the recovery scheme is using a log with immediate updates, show the state of the log if the system crashes just after “\( \text{Read}(B) \)”. Assume all possible log entries have been written to disk. Explain how the system recovers.