CIS 560
Database System Concepts

Hour Exam 1 (Closed-Book, Open-Notes, Open-Mind)
Friday, 06 October 2006

Instructions and Notes

- You are permitted two (2) double-sided or four (4) single-sided, typewritten or handwritten pages of notes.
- No calculators or computing devices are needed or permitted on this exam.

- You should have a total of 8 pages; write your name on each page.
- There are six (6) problems. You have 75 minutes for this exam. Budget your time carefully.

- In the interest of fairness to all students, no questions shall be answered during the test concerning definitions.
- If you believe there is an error or ambiguity in any question, notify the instructor and state your assumptions.

- Your answers on short answer and essay problems shall be graded for originality as well as for accuracy.
- Use the space provided for your answers; you may add additional pages if needed.
- Select exactly one answer for each true/false and multiple choice question.
- Show your work on problems and proofs.
- There are a total of 100 possible points in this exam.

Instructor Use Only

1. _____ / 15
2. _____ / 20
3. _____ / 10
4. _____ / 15
5. _____ / 25
6. _____ / 15

Total _____ / 100
Consider the following schema for problems 1-5:

- **Stations**: (station-name, area-of-study)
- **Infiltrators**: (name, area-of-study, status)
- **Personnel**: (first-name, surname, station-name, year)
- **Incidents**: (station-name, year)

### Stations

<table>
<thead>
<tr>
<th>station-name</th>
<th>area-of-study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swan</td>
<td>Electromagnetism</td>
</tr>
<tr>
<td>Pearl</td>
<td>Psychology</td>
</tr>
<tr>
<td>Arrow</td>
<td>Unknown</td>
</tr>
<tr>
<td>Staff</td>
<td>Medical</td>
</tr>
<tr>
<td>Hydra</td>
<td>Zoology</td>
</tr>
<tr>
<td>Flame</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

### Infiltrators

<table>
<thead>
<tr>
<th>name</th>
<th>area-of-study</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ben</td>
<td>Psychology</td>
<td>Living</td>
</tr>
<tr>
<td>Ethan Rom</td>
<td>Tactical</td>
<td>Deceased</td>
</tr>
<tr>
<td>Goodwin</td>
<td>Psychology</td>
<td>Deceased</td>
</tr>
</tbody>
</table>

### Personnel

<table>
<thead>
<tr>
<th>first-name</th>
<th>surname</th>
<th>station-name</th>
<th>year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom</td>
<td>Friendly</td>
<td>Staff</td>
<td>1975</td>
</tr>
<tr>
<td>Ben</td>
<td>NULL</td>
<td>Swan</td>
<td>1971</td>
</tr>
<tr>
<td>Ethan Rom</td>
<td>Rom</td>
<td>Door</td>
<td>1985</td>
</tr>
<tr>
<td>NULL</td>
<td>Goodwin</td>
<td>Door</td>
<td>1992</td>
</tr>
<tr>
<td>Juliet</td>
<td>NULL</td>
<td>Hydra</td>
<td>1982</td>
</tr>
<tr>
<td>Bea</td>
<td>Klugh</td>
<td>Door</td>
<td>1986</td>
</tr>
</tbody>
</table>

### Incidents

<table>
<thead>
<tr>
<th>station-name</th>
<th>year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swan</td>
<td>2004</td>
</tr>
<tr>
<td>Staff</td>
<td>NULL</td>
</tr>
<tr>
<td>Hydra</td>
<td>2004</td>
</tr>
<tr>
<td>Flame</td>
<td>1981</td>
</tr>
</tbody>
</table>
1. **Relational Algebra (3 parts, 5 points each).** Convert the following queries from English into relational algebra or vice versa (that is, write the expression or explain what it means; don't just evaluate them). You may use the natural join operator, \( \bowtie \). All attributes have string type except year in Personnel and Incidents, which are integers.

a) A relation containing the name of each Infiltrator whose area-of-study is “Psychology”, along with each corresponding station-name from Stations.

b) \[ \pi_{\text{first-name}, \text{surname}} \left( \sigma_{\text{year-joined} \leq \text{year}} \left( \rho_{\text{first-name}, \text{last-name}, \text{station-name}, \text{year-joined}} (\text{Personnel}) \times \text{Incidents}) \right) \right) \]

c) What result does the query in part (b) return? (Hint: a NULL field will fail a comparison or Theta join in this case.)
2. SQL (4 parts. 5 points each).

Recall that

\[
\text{select } A_1, A_2, \ldots, A_n \\
\text{from } r_1, r_2, \ldots, r_m \\
\text{where } P
\]

is equivalent to the relational algebra expression

\[
\Pi_{A_1, A_2, \ldots, A_n}(\sigma_P(r_1 \times r_2 \times \ldots \times r_m))
\]

**SELECT query. Give SQL expressions returning:**

a) A relation containing the name of each Infiltrator whose area-of-study is “Psychology”, along with each corresponding station-name from Stations. (Do not assume that the tables Infiltrator or Stations contain distinct values.)

b) A relation containing the year in which each person working at the “Medical” station joined.

What would be returned by:

c) a left outer join between Personnel and Incidents?

d) a full outer join between Personnel and Incidents?
3. Domain Relational Calculus and Query-by-Example (2 parts, 5 points each). Repeat Problem 2, parts a and c, using domain relational calculus and Query-by-Example (QBE).

a) (Specify using domain relational calculus) The first-name and surname of anyone working at station “Staff”. Show what is returned.

b) (Specify using QBE) The average year in which anyone at station “Door” joined. Show what is returned.
4. More Relational Algebra (3 parts, 5 points each).

    Suppose we wanted years in which an incident happened at all stations in a list (stored as a table).
    Give two ways of getting this answer: one with relational division (\(\div\)), and one without.

    First method (with \(\div\)):

    Second method (without \(\div\)):

    Result:
    What quotient is returned for dividend \textbf{Incidents} and divisor station list [Swan, Hydra]?


a) (5 points) In problem 3(a), you found the first-name and surname of anyone working at station "Staff". Is either one a candidate key for Personnel? Why or why not?

b) (15 points) Suppose the Dharma Initiative hired you to design an entity-relational data model for their enterprise database. Define a relationship set Stationed-At between Personnel and Stations. Draw how it relates them. Be sure to draw the attributes of each entity in E-R notation. Indicate which relationships are many-to-one, one-to-one, or one-to-many.

Once you understand database normalization, you should see why the above entity design needs improvement.

c) (5 points) Ben comes to you and asks for a head count of living people who can be called up from each station. What's wrong with Stations that prevents you from doing a proper COUNT? State your answer in terms of participation in a relationship set or what isn't (mathematically) one-to-one/onto.

You are hired as a contractor to design the lending database for the Widmore University Libraries. The entities include **Library** (subject areas: Law, Medicine, Engineering, Main), **Book**, and **Borrower**.

Draw an E-R diagram for a schema that includes the above entities, typical attributes, and a ternary (3-way) relationship set **Loan**, with identifying set **Library**. Show the cardinality of relationships and draw at least three attributes belonging to **Loan**. Explain why they belong to **Loan** and not to any of the entity sets.