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

station-name	area-of-study
Swan	Electromagnetism
Pearl	Psychology
Arrow	Unknown
Staff	Medical
Hydra	Zoology
Flame	Unknown

Infiltrators

name	area-of-study	status
Ben	Psychology	Living
Ethan Rom	Tactical	Deceased
Goodwin	Psychology	Deceased

Personnel

first-name	surname	station-name	year
Tom	Friendly	Staff	1975
Ben	NULL	Swan	1971
Ethan	Rom	Door	1985
NULL	Goodwin	Door	1992
Juliet	NULL	Hydra	1982
Bea	Klugh	Door	1986

Incidents

station-name	year
Swan	2004
Staff	NULL
Hydra	2004
Flame	1981

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) $\pi_{\text{first-name, surname}}$
$(\sigma_{\text{year-joined} \le \text{year}} (\rho_{\text{S(first-name, last-name, station-name, year-joined)}} (\textbf{Personnel}) \times \textbf{Incidents}))$

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

```
select A1, A2, ..., An
from r1, r2, ..., rm
where P
```

is equivalent to the relational algebra expression $\prod_{A_1, A_2, \ldots, A_n}(\sigma_P(r_1 \times r_2 \times \ldots \times r_m))$

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 <u>not</u> 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 <u>two</u> ways of getting this answer: one with relational division (÷), and one without.

First method (with ÷):

Second method (without ÷):

Result:
What quotient is returned for dividend **Incidents** and divisor station list [Swan, Hydra]?

5. Keys and E-R Diagram Design (3 parts, 25 points total).

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.

6. More E-R Diagram Design (15 points).

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.