CMSC 424–0101 Fall 2002 **Midterm Exam (60 points, 60 minutes)** Thu, 24 Oct 2002

Name: _____

This midterm is **open book, open notes**, but there can be **no sharing** of any material. Some questions in this midterm use the database schema and sample instance depicted below (which we have encountered earlier, in the quiz). For brevity, relational algebra expressions abbreviate the schemas of the tables as indicated next to the table names below. The last page of this midterm duplicates the tables below. You may tear it off for easier reference as you work on the questions. You do not need to reattach it.

DigiCams $D(M, L, P, A)$

Manufacturer varchar(30)	Model varchar(30)	MPix real	MaxAp real
Canon	G3	3.9	2.0
Canon	G2	3.9	2.0
Nikon	4500	3.87	2.6
Olympus	4040	3.9	1.8
Olympus	3030	3.14	2.8

Catalog $C(S, M, L, P)$

Store varchar(30)	Manufacturer varchar(30)	Model varchar(30)	Price real
Joe's Place	Olympus	4040	599.99
Snap Chap	Olympus	4040	655.00
Zonama	Canon	G2	698.50
Zonama	Olympus	3030	488.55
Zonama	Nikon	4500	589.95

Reviews $R(S, T, M, L, R, D)$

Source varchar(30)	RDate date	Manufacturer varchar(30)	Model varchar(30)	Rating integer	Desc CLOB
Photo Life	2002-02-03	Canon	G2	9	Our lab...
Photo Life	2001-08-22	Olympus	4040	7	When I...
PC Mag	2002-09-15	Nikon	4500	8	Often a...

1. (1 pt) Write your name in the space provided above.

2. (4×5 pts) Write queries as directed, referring to the database introduced earlier.

   (a) Write a SQL query that lists cameras (manufacturer and model for each) that have been reviewed by a source whose name contains the string `Photo` and does not contain the string `%defunct`. (Note the literal % sign.)

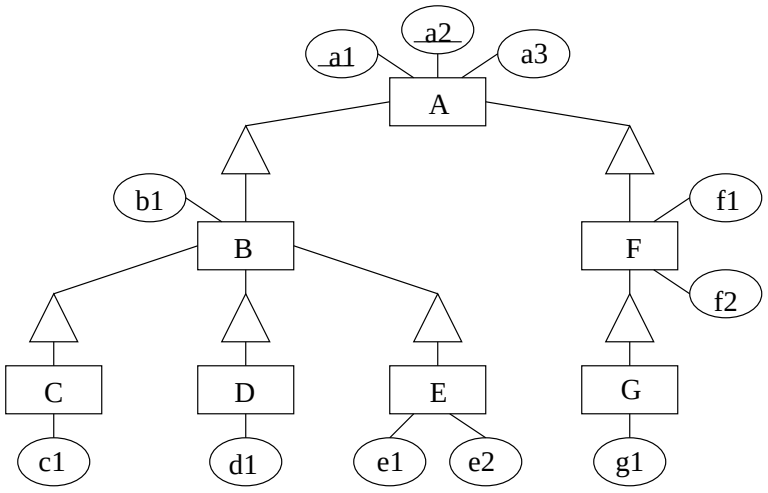
(b) Write a SQL query that is equivalent to the natural join of the three tables (using the long attribute names, not the abbreviations).

(c) Write a SQL query for pairs of cameras (manufacturers and models) that have the same sensor resolution (`MPix` value). Do not pair a camera with itself, and include each qualifying pair (not ordered pair) only once.

(d) Write an extended algebra query equivalent to the following SQL query (which you may recognize from the quiz):

```
select D.Manufacturer, D.Model, min(C.Price)
from DigiCams D, Catalog C
where D.Manufacturer = C.Manufacturer and D.Model = C.Model
group by D.Manufacturer, D.Model
order by D.Manufacturer asc, D.Model desc;
```

3. (9 pts) Convert the following ER diagram in to a relational schema using each of the three methods discussed in class. (Write down the three relational schemas; you do not need to explain your work.) Note that a_1 and a_2 are the only key attributes.



4. (10 pts) Given that the instance of `DigiCams` depicted earlier is valid, list five functional dependencies that **cannot** hold. Justify your answer **briefly**. For this question, assume that no key attributes are specified for `DigiCams` (i.e., ignore the implications of the underlining of column names.) Hint: Think about how you could tell if an instance does not satisfy a functional dependency.

5. (3×5 pts)

   In a database system that permits duplicates in tables (e.g., Oracle, PostgreSQL), there are two methods for representing bags: The first method, which we shall call the *expanded representation*, simply duplicates tuples as needed. For example, a bag with two instances of $(10, 20)$ and one instance of $(30, 20)$ is represented in the table as $((10, 20), (10, 20), (30, 20))$. The second method, which we call the *concise representation*, appends a count attribute to the schema to encode the number of instances of each tuple in the bag. Our earlier example is represented as $\{(10, 20, 2), (30, 20, 1)\}$ in concise form. (This representation is similar to the one discussed in class. The only difference is that we flatten the tuples, i.e., we write $(10, 20, 2)$ instead of $((10, 20), 2)$.) To make a concise representation B unique, we impose two constraints: (1) The counts must be strictly positive. That is, if $(a_1, \ldots, a_k, n) \in B$, then $n > 0$. (2) Tuples are uniquely identified by their non-count attributes. That is, if $(a_1, \ldots, a_k, n_1) \in B$ and $(a_1, \ldots, a_k, n_2) \in B$ then $n_1 = n_2$.

   (a) Given a table `R(A, B, C)` that encodes a bag using the expanded representation, and another table `CR(A, B, C, N)` that is empty, write a SQL statement to populate `CR` with the concise representation of the bag in `R`. (Assume all attributes

are of integer type.)

(b) Given tables `CR(A, B, C, N)` and `CS(A, B, C, N)` that store bags in the concise representation (meaning the `N` attributes in both schemas represent counts), write a SQL query for the **concise representation** of the bag union of `CR` and `CS`.

(c) Given the tables in Question 5b, write a SQL query for the **concise representation** of the result of the bag algebra query $\pi^{\mathcal{B}}_{A,B,C,B2,C2}(CR \bowtie^{\mathcal{B}}_{C=A2} \rho_{T(A2,B2,C2)} CS)$. (Remember to interpret CR and CS as concise representations of bags.)

6. (5 pts) You may think of this question as *ER Lego*. You are asked to build an ER diagram out of the given components: You have three boxes, four diamonds, five ovals, and an unlimited number of connecting lines (including arrows) of all sorts. Each of the boxes and diamonds may be used in either a single-line or a double-line mode. Similarly, each oval may be used in either underlined or non underlined mode. Your goal is to maximize the size (number of attributes) of the largest key. You do not have to use all pieces.

   Depict your ER diagram below. Indicate the size of the largest key (and an entity set with such a key). You do not need to justify your answer. Hint: Avoid enumerating all possibilities!

Scratch Page

You may tear off this page for easier reference. **This page will not be graded.**

DigiCams $D(M, L, P, A)$

Manufacturer varchar(30)	Model varchar(30)	MPix real	MaxAp real
Canon	G3	3.9	2.0
Canon	G2	3.9	2.0
Nikon	4500	3.87	2.6
Olympus	4040	3.9	1.8
Olympus	3030	3.14	2.8

Catalog $C(S, M, L, P)$

Store varchar(30)	Manufacturer varchar(30)	Model varchar(30)	Price real
Joe's Place	Olympus	4040	599.99
Snap Chap	Olympus	4040	655.00
Zonama	Canon	G2	698.50
Zonama	Olympus	3030	488.55
Zonama	Nikon	4500	589.95

Reviews $R(S, T, M, L, R, D)$

Source varchar(30)	RDate date	Manufacturer varchar(30)	Model varchar(30)	Rating integer	Desc CLOB
Photo Life	2002-02-03	Canon	G2	9	Our lab...
Photo Life	2001-08-22	Olympus	4040	7	When I...
PC Mag	2002-09-15	Nikon	4500	8	Often a...