UNIVERSITY OF TORONTO
Faculty of Arts and Science

DECEMBER EXAMINATIONS 2006
CSC343H|F

Duration -- 3 hours

No aids allowed

This exam is worth 35% of your final mark. Please answer all questions in the space provided. You may use the back of pages for rough work. If you'd like us to grade something there, please indicate so clearly on the front of the page. Good luck!

Name_____
(Please underline last name)

Student number_____

Q1._____**/70**

Q2._____**/70**

Q3._____**/80**

Q4._____**/50**

Q5._____**/80**

TOTAL_____**/350**

Question 1 [Short questions; 70 marks] There are 14 T/F questions below. A correct answer is worth 5 marks; a wrong answer is worth -4 marks; no answer is worth 0 marks. The minimum mark you can receive for this problem is 0/70.

(a) ____ In Datalog, one can define queries that can ***not*** be expressed in the Relational Algebra.

(b) ____ Assuming that `emp(emp#,addr,sal)` is a database relation, the following Datalog rule is unsafe

```
raise(emp#,newsal) :- newsal=1.1*sal,emp(emp#,addr,sal)
```

(c) ____ Assume that relationship `WorksFor` relates entities `Worker`, `Project` and has two attributes `from`, `to` that specify the period a worker worked for a project. According to this E-R schema, a worker can't work for the same project more than once.

(d) ____ In the E-R Model, a weak entity e_1 is deleted whenever another entity e_2 is deleted and e_2 is part of e_1's key.

(e) ____ There is a functional dependency from postal code to address -- i.e., `postalCode` → `address` -- according to Canada Post rules.

(f) ____ If A, B, C, D are single attributes, and $F = \{AB \rightarrow CD\}$, then $F = F^+$

(g) ____ The decomposition of relation `Emp(emp#,addr,city)` into relations R_1`(emp#,addr)` and R_2`(addr,city)` is lossless.

(h) ____ There exist relational schemas that are in 3NF but not in BCNF.

(i) ____ A DTD limits the labels that can be used in an XML document, but not the order in which they can appear.

(j) ____ In Static SQL, a transaction can be terminated before the end of a program by statements EXEC SQL COMMIT or EXEC SQL ROLLBACK.

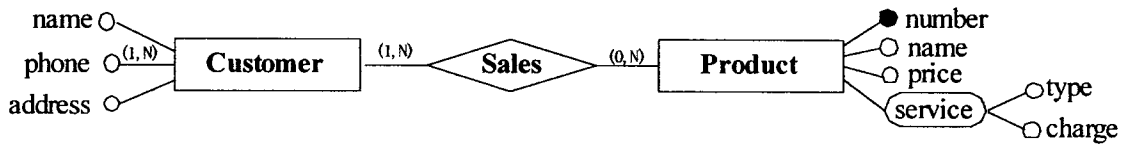
(k) ____ $\pi_A(R - S) = \pi_A(R) - \pi_A(S)$ is always true.

(l) ____ Recursive queries can be expressed in SQL by using recursive views.

(m) ____ In the schema R[ABCDE] with FDs A → BC, D → E, AE is a key.

Question 2 [Entity-Relationship model; 70 marks] *Amy's Bikes* is a new bike shop located in a Toronto suburb, offering a wide range of bicycles and related accessories. Amy, the owner, has been conducting her daily business mostly on paper. She records sales on preprinted forms that contain the invoice number and date of the sale, the customer and the employee involved in the sale and the product being sold. Employee and customer information is maintained on sheets of paper. For each employee, this includes his/her social insurance number, first and last name, and home phone number. For each customer, Amy records the first and last name, and at least one phone number and home address (consisting of street number and post code). Different customers may have the same name. To keep track of the product inventory, Amy uses a spreadsheet program to record the number, name, price and quantity of the products in stock. For each product, a range of after-sales services is offered. The spreadsheet program is also used to list the type (e.g., repair, exchange) and charge for each service. A product may have any number of services associated with it, but a service can only be associated with a single product.

Amy spends a lot of time maintaining this information. Recently, she has decided to use database technology (!) to manage all this data. After a brief study of database design techniques, Amy drew her very first ER diagram, shown below. For simplicity, (1,1) cardinality constraints are omitted everywhere in the diagram:



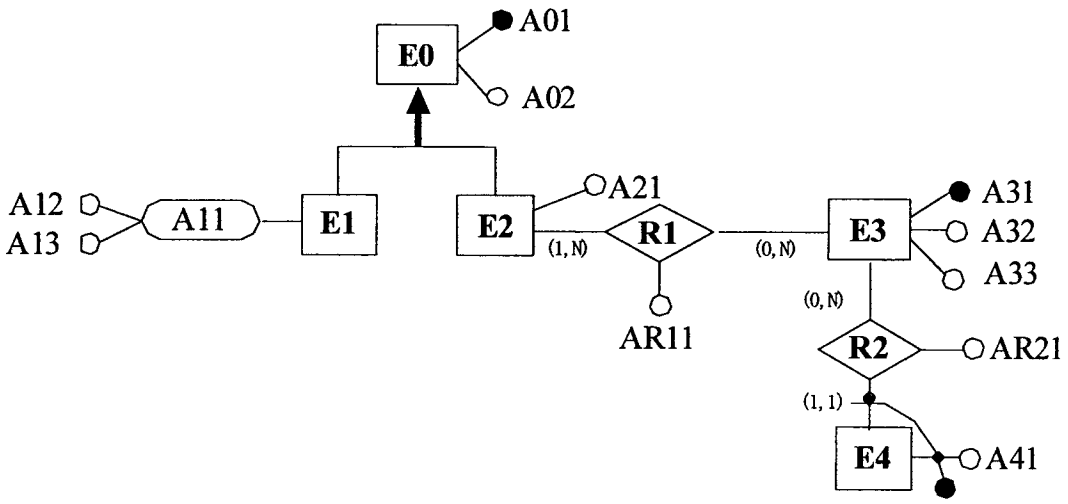
(2.a) [10 marks] Unfortunately, Amy's E-R schema is both incomplete and incorrect with respect to the requirements stated above. To help her, give one missing attribute for Sales and explain why Service is modeled inappropriately.

Missing attribute:

Error for Service***:***

(2.b) [30 marks] Now let's help Amy to design an E-R schema that better captures the requirements. The schema should contain all relevant entities and relationships mentioned in the requirements. You can omit (1,1) cardinality constraints. All other integrity constrains have to be represented explicitly using proper E-R notation. If you think there is ambiguity in the requirements, make your assumptions and state them clearly.

(2.c) [30 marks] Convert the follow E-R schema to a relational one, using suitable transformations. All missing cardinalities can be assumed to be (1,1).



Question 3 [XML query processing; 80 marks] Figure 3.1 shows a Document Type Definition (DTD) file, dcs.*dtd,* describing some information about an academic department.

```
<!ELEMENT DCS (Professors, Students, Courses, Lab+)>
<!ELEMENT Professors (Professor+)>
<!ELEMENT Students (Student*)>
<!ELEMENT Courses (Course+)>
<!ELEMENT Professor (ProfName, Email)>
<!ELEMENT Student (StudName, Level, CrsTaken*, CrsTaking*, Supervisor?)>
<!ELEMENT Course (Title, OfferedYear, Instructor)>
<!ELEMENT ProfName (#PCDATA)>
<!ELEMENT Email (#PCDATA)>
<!ELEMENT StudName (#PCDATA)>
<!ELEMENT Level (#PCDATA)>
<!ELEMENT CrsTaken (Grade)>
<!ELEMENT CrsTaking EMPTY>
<!ELEMENT Supervisor (#PCDATA)>
<!ELEMENT Title (#PCDATA)>
<!ELEMENT OfferedYear (#PCDATA)>
<!ELEMENT Instructor (#PCDATA)>
<!ELEMENT Grade (#PCDATA)>
<!ELEMENT Lab (#PCDATA)>
<!ATTLIST Professor pid ID #REQUIRED>
<!ATTLIST Student snum ID #REQUIRED>
<!ATTLIST Course cnum ID #REQUIRED>
<!ATTLIST CrsTaken cnum IDREF #REQUIRED>
<!ATTLIST CrsTaking cnum IDREF #REQUIRED>
```

Figure 3.1: dcs.dtd

An XML document, *dcs.xml*, is given in Figure 3.2 (page 9). Please answer the following questions with respect to dcs.dtd and dcs.xml.

(3.a) [10 marks] Is dcs.xml a valid XML document wrt dcs.dtd? If it is not, please explain what modification(s) should be done ***on the DTD file*** to make the XML file valid.

(3.b) [10 marks] Briefly explain what the following XPath expression returns:

```
//Students[Student/@snum="s2002"].
```

(3.c) [10 marks] For the following XQuery expression, write down the output of the query against dcs.xml.

```
<result>
{
for $p in doc("dcs.xml")/DCS/Professors/Professor
where $p/@pid="p1001"
return $p/ProfName
}
</result>
```

(3.d) [20 marks] For the following XQuery expression, write down the output of the query when evaluated against dcs.xml.

```
<result>
{
for $s in doc("dcs.xml")//Student,
    $c in doc("dcs.xml")//Course
where some $sc in $s/CrsTaking satisfies
                $sc/@cnum=$c/@cnum and
                $c/Instructor="Marley"
return <Student><SNum>$s/@snum</SNum>
        <SName>$s/StudName</SName>
        <SLevel>$s/Level</SLevel>
        </Student>
}
</result>
```

(3.e) [30 marks] Define a query in XQuery that lists the names of all of Prof. Olay's students (i.e., students whose supervisor is Prof. Olay). These names are to be grouped under Prof. Olay's name. For example, the output should look as follows when the query is executed against dcs.xml:

```
<Professor>
        <ProfName>Olay</ProfName>
        <Students>
                <Student>
                        <SName>Adam</SName>
                </Student>
                <Student>
                        <SName>Eva</SName>
                </Student>
        </Students>
</Professor>
```

```
<DCS>
    <Professors>
        <Professor pid="p1001">
            <ProfName>Olay</ProfName>
            <Email>olay@university.edu</Email>
        </Professor>
        <Professor pid="p1002">
            <ProfName>Marley</ProfName>
            <Email>marley@university.edu</Email>
        </Professor>
    </Professors>
    <Students>
        <Student snum="s2001">
            <StudName>Joe</StudName>
            <Level>undergrad</Level>
            <CrsTaking cnum="c3002"/>
        </Student>
        <Student snum="s2002">
            <StudName>Adam</StudName>
            <Level>grad</Level>
            <CrsTaken cnum="c3001">
                <Grade>80</Grade>
            </CrsTaken>
            <Supervisor>Olay</Supervisor>
        </Student>
        <Student snum="s2003">
            <StudName>Eva</StudName>
            <Level>grad</Level>
            <Supervisor>Olay</Supervisor>
        </Student>
    </Students>
    <Courses>
        <Course cnum="c3001">
            <Title>Databases</Title>
            <OfferedYear>2005</OfferedYear>
            <Instructor>Olay</Instructor>
        </Course>
        <Course cnum="c3002">
            <Title>Programming</Title>
            <OfferedYear>2006</OfferedYear>
            <Instructor>Marley</Instructor>
        </Course>
    </Courses>
</DCS>
```

Figure 3.2: dcs.xml

Question 4 [More SQL and Relational Algebra; 50 marks] Consider the following schema and the corresponding database:

Sailors (*sid: integer, sname: string, rating: integer, age: real*)
Boats (*bid: integer, bname: string, color: string*)
Reserves (*sid: integer, bid: integer, day: date*)

The key attributes of each relation are underlined.

Sailors

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

Reserves

sid	bid	day
22	101	10/10/98
22	102	10/10/98
22	103	10/8/98
22	104	10/7/98
31	102	11/10/98
31	103	11/6/98
31	104	11/12/98
64	101	9/5/98
64	102	9/8/98
74	103	9/8/98

Boats

bid	bname	color
101	Intelake	blue
102	Interlake	red
103	Clipper	green
104	Marine	red

(4.a) [10 marks] Show the result of the following query evaluated wrt the given database:

```
SELECT    S.sid, S.sname
FROM      Sailors S
WHERE     S.rating > ALL (   SELECT    S2.rating
                             FROM      Sailors S2
                             WHERE     S2.sname = 'Horatio' )
```

(4.b) [20 marks] Write the following query ***in both*** Relational Algebra and SQL:

Find the name of sailors who have never reserved a boat named 'Interlake'.

(4.c) [20 marks] Write the following query in SQL:

Find the bid and the number of reservations for each 'red' boat with at least two reservations.

Question 5 [Functional Dependencies; 80 marks] Given R = (*R; F*) where *R* = *ABCDEGH* and *F* = *{ABH → C, A → DE, BGH → C, C → ADH, BH → GE}*.

(5.a) [20 marks] Compute F^+. Since there are many FDs in F^+, you may want to compute F^+ by computing X^+_F only for sets of attributes X such that $X \neq X^+_F$.

(5.b) [30 marks] Compute the minimal cover of F.

(5.c) [30 marks] Place R in 3NF.

[End of exam]