

Name: _____

This quiz is **open book, open notes**, but there can be no sharing of any material. You can use the Internet, but only as a library. If you are not sure if something is allowed, check with me.

Many questions in this quiz use the database schema and sample instance depicted below. (*This database is identical to the one used in the midterm.*) The **Accesses** table is essentially a log of HTTP accesses. For simplicity, we assume a single-threaded HTTP server that services at most one request at any time, so that the timestamp is sufficient to identify a request. For each request, the table records the originating IP address, the timestamp, the command, and a status code. The **SourceMap** table stores information about the association between IP addresses and users at different times. A tuple (a, b, e, u, p) in this table indicates that a IP address a was used by user u from time b to time e (continuously); p is a priority code that is used to indicate the importance of each tuple. The type of each attribute appears directly below its name. *Primary key attributes* are underlined.

When asked for queries, you must provide answers that work for all possible database instances, not just the example instance depicted below.

The following tables are repeated on the last page of the quiz. You may detach that page and use it for reference. There is no need to reattach it

Accesses

<u>FAddr</u> varchar(15)	<u>ATime</u> timestamp	Cmd varchar(100)	SCode integer
198.137.240.92	2003-01-30 14:01:02	GET / HTTP/1.1	200
160.111.252.106	2002-10-10 10:10:23	GET /robots.txt /HTTP/1.1	404
160.111.252.106	2002-10-10 10:11:36	HEAD /GETME.txt /HTTP/1.1	404
198.137.240.92	2003-01-30 14:10:32	HEAD / HTTP/1.0	200
198.137.240.92	2003-01-30 14:10:55	GET /secret.txt HTTP/1.0	200

SourceMap

<u>IPAddr</u> varchar(15)	<u>UseBegin</u> timestamp	UseEnd timestamp	User varchar(100)	PCode integer
198.137.240.92	2003-01-30 11:00:32	2003-01-30 15:03:46	H. White	1
198.137.240.92	2002-02-15 8:00:02	2002-03-30 18:13:26	I.S. Abel	2
160.111.252.106	2002-02-25 9:00:41	2002-02-30 9:13:26	S. Smith	11

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

5. (6 pts.) Suppose we augment our database with a table `SourceMapToDo` with columns `IPAddr` and `ATime` of types `varchar(15)` and `timestamp`, respectively. Write a trigger to achieve the following behavior: Whenever an insertion to `Accesses` introduces an IP address (in the `FAddr` attribute) that does not occur in the `SourceMap` table, that IP address and the access time are inserted into the `SourceMapToDo` table.
6. (6 pts.) Suppose we have defined a class `TInt` as follows, for representing time intervals in our database. (We assume `timestamp` is an ODL datatype analogous to the SQL counterpart.)

```
class TInt (extent TInts key (tBegin, tEnd) {
    attribute timestamp tBegin;
    attribute timestamp tEnd;
};
```

Present an ODL declaration of a class `SourceMap` that may be used to store the information stored by the `SourceMap` table. We wish to use `TInt` objects for storing time intervals. The rest of the information must reside in the `SourceMap` objects. Use attribute and relationship names that are similar to those in the `SourceMap` table. Indicate any modifications that may be needed to the definition of `TInt`. Declare keys and extents if you can. Hint: Pay attention to the multiplicities of relationships.

7. (6 pts.) Using the definition of `TInt` from Question 6, write an OQL query that computes the average length of intervals. (You may assume that arithmetic operators behave appropriately for the timestamp datatype.)

8. (7 pts.) [This question is probably the hardest one on this test.] For the extent for class `SourceMap` in your answer to Question 6, write an OQL query that returns a list of the average *connect time* for each `pCode` and `ipAddr`, excluding `pCode-ipAddr` pairs for which the total connect time is less than one hour. We define connect time to be the length of the time interval denoting the connection.

The result should be a list of structs of the form `(code: c, ip: i, avgLen: l)`, indicating that the average connect time for `pCode` `c` and `IPAddr` `i` is `l`. This list should be sorted in descending order of `avgLen` (primary order), ascending order of `code` (secondary order), and ascending order of `ip` (ternary order).

In standard SQL, if `a` and `b` are of type *timestamp* then `a - b` is of type *interval*. A literal time interval of length one hour may be specified as “`interval '01:00:00' hour to minute`.” We will use these conventions for OQL too.

9. (6 pts.) We say that user u had a *close encounter* with user v if there is at least one IP address i such that v started using i less than an hour after u stopped using i . We say u and v are *related* if there is a chain of close encounters from u to v . In other words, related is the transitive closure of close encounter. (Note that related is neither symmetric nor reflexive: u related to v does not imply v related to u and u is not related to u .)

Write a Datalog program for *related*. Remember to ensure that your program is safe and stratified. Assume that subtracting two timestamps yields a time interval and that a literal interval of length one hour may be written as “1h.”

Scratch page

Material here will not be graded. You may detach and discard this page.

Accesses

<u>FAddr</u> varchar(15)	<u>ATime</u> timestamp	<u>Cmd</u> varchar(100)	<u>SCode</u> integer
198.137.240.92	2003-01-30 14:01:02	GET / HTTP/1.1	200
160.111.252.106	2002-10-10 10:10:23	GET /robots.txt /HTTP/1.1	404
160.111.252.106	2002-10-10 10:11:36	HEAD /GETME.txt /HTTP/1.1	404
198.137.240.92	2003-01-30 14:10:32	HEAD / HTTP/1.0	200
198.137.240.92	2003-01-30 14:10:55	GET /secret.txt HTTP/1.0	200

SourceMap

<u>IPAddr</u> varchar(15)	<u>UseBegin</u> timestamp	<u>UseEnd</u> timestamp	<u>User</u> varchar(100)	<u>PCode</u> integer
198.137.240.92	2003-01-30 11:00:32	2003-01-30 15:03:46	H. White	1
198.137.240.92	2002-02-15 8:00:02	2002-03-30 18:13:26	I.S. Abel	2
160.111.252.106	2002-02-25 9:00:41	2002-02-30 9:13:26	S. Smith	11