Name: ________________________________

This exam 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 exam use the database schema and sample instance depicted below. (This database is identical to the one used in the midterm and second quiz.) 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. In relational algebra and Datalog, you may abbreviate the schemas for the Accesses and SourceMap tables as \(A(F, A, C, S)\) and \(S(I, B, E, U, P)\), respectively.

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 exam. You may detach that page and use it for reference. There is no need to reattach it

\[
\begin{array}{|c|c|c|c|}
\hline
\text{FAddr} & \text{ATime} & \text{Cmd} & \text{SCode} \\
\text{varchar(15)} & \text{timestamp} & \text{varchar(100)} & \text{integer} \\
\hline
198.137.240.92 & 2003-01-30 14:01:02 & GET / HTTP/1.1 & 200 \\
160.111.252.106 & 2002-10-10 10:10:23 & GET /robots.txt /HTTP/1.1 & 404 \\
160.111.252.106 & 2002-10-10 10:11:36 & HEAD /GETME.txt /HTTP/1.1 & 404 \\
198.137.240.92 & 2003-01-30 14:10:32 & HEAD / HTTP/1.0 & 200 \\
198.137.240.92 & 2003-01-30 14:10:55 & GET /secret.txt HTTP/1.0 & 200 \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{IPAddr} & \text{UseBegin} & \text{UseEnd} & \text{User} & \text{PCode} \\
\text{varchar(15)} & \text{timestamp} & \text{timestamp} & \text{varchar(100)} & \text{integer} \\
\hline
198.137.240.92 & 2003-01-30 11:00:32 & 2003-01-30 15:03:46 & H. White & 1 \\
198.137.240.92 & 2002-02-15 8:00:02 & 2002-03-30 18:13:26 & I.S. Abel & 2 \\
\hline
\end{array}
\]

1. (1 pt.) Write your name in the space provided above.
2. (8 pts.) Write a SQL query that returns the names of users who issued a command with SCode greater than 300.

3. (8 pts.) Write a relational algebra query that is equivalent to the query in Question 2.

4. (8 pts.) Write a Datalog query that is equivalent to the query in Question 2.

5. (8 pts.) Consider the following relation (schema) and functional dependencies. (A copy of these appears on the last page of the exam, which you may detach.)

\[ R(A, B, C, D, E, F) \]

\[ f_1: \ ABC \rightarrow D \]
\[ f_2: \ CE \rightarrow B \]
\[ f_3: \ E \rightarrow C \]
\[ f_4: \ BF \rightarrow AC \]
\[ f_5: \ E \rightarrow F \]
\[ f_6: \ F \rightarrow E \]

Compute \( \{A, B\}^+, \{C, D\}^+, \) and \( \{E, F\}^+. \)
6. (8 pts.) List all keys for the schema of Question 5.

7. (8 pts.) List all BCNF violations for the schema of Question 5.

8. (8 pts.) Normalize the schema of Question 5 to BCNF. Justify your answer briefly. (At each step, note the projected functional dependencies, keys, and dependencies, if any, used to decompose the schema.)
9. (9 pts.) Mark all conflicting operations in the following schedule using the arrow notation discussed in class.

\[ r_1(x), r_2(x), r_2(y), w_1(y), r_3(y), w_1(x), r_3(x), w_3(y) \]

10. (9 pts.) Is the schedule of Question 9 serializable? Justify your answer briefly.
11. (9 pts.) Given an example of a schedule that is permitted by SQL’s *read committed* isolation level but not permitted by the *serializable* isolation level. Use the notation used in class and justify your answer.

12. (9 pts.) We say $v$ used $i$ immediately following $u$, denoted by $nextUser(i, u, v)$, if and only if there exist timestamps $t_u$ and $t_v$ such that $u$ stopped using $i$ at $t_u$, $v$ started using $i$ at $t_v$, and there is no other use of $i$ (by any user) between $t_u$ and $t_v$. Write a Datalog program for $nextUser$. Hint: Remember to verify that your Datalog programs safe and stratified.

13. (9 pts.) We say that a user $u$ is a *secret reader*, denoted by $secretReader(u)$ if and only if $u$ issued a command “GET /secret.txt HTTP/1.0.” Similarly, we say a user $u$ is a *junk reader*, denoted by $junkReader(u)$ if and only if $u$ issued a command “GET /junk.txt HTTP/1.0.” Write a Datalog program for the predicates $secretReader$ and $junkReader$.
14. (9 pts.) Let the predicates nextUser, secretReader, and junkReader be as defined in Questions 12 and 13. We say there is chain of use from u to v, denoted by useChain(u, v), if and only if there exist, for \( k \geq 2 \), users \( x_1, x_2, \ldots, x_k \) such that \( u = x_1, v = x_k, \forall i \in [1, k - 1] \exists a : nextUser(a, x_i, x_{i+1}) \). We say there is an s-j chain from u to v, denoted by sjChain(u, v), if there is a chain of use from u to v in which users alternate between secret readers and junk readers. Write a Datalog program for sjChain.

15. (9 pts.) Write a SQL-99 query that is equivalent to the Datalog query in Question 14. You may use views. [Hint: The required query is rather long, but not really difficult if you use the Datalog version as a guide. Proceed step-by-step.]
Score Page
Do not write on this page. It will not be graded.
Do NOT detach this page.

Score Summary

<table>
<thead>
<tr>
<th>Question</th>
<th>Score</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>14</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>120</strong></td>
<td></td>
</tr>
</tbody>
</table>

Do not write on this page. It will not be graded.
Do NOT detach this page.
# Scratch Page

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

<table>
<thead>
<tr>
<th>IPAddr</th>
<th>UseBegin</th>
<th>UseEnd</th>
<th>User</th>
<th>PCode</th>
</tr>
</thead>
<tbody>
<tr>
<td>198.137.240.92</td>
<td>2003-01-30 11:00:32</td>
<td>2003-01-30 15:03:46</td>
<td>H. White</td>
<td>1</td>
</tr>
<tr>
<td>198.137.240.92</td>
<td>2002-02-15 08:00:02</td>
<td>2002-03-30 18:13:26</td>
<td>I.S. Abel</td>
<td>2</td>
</tr>
<tr>
<td>160.111.252.106</td>
<td>2002-02-25 09:00:41</td>
<td>2002-02-30 09:13:26</td>
<td>S. Smith</td>
<td>11</td>
</tr>
</tbody>
</table>

$R(A, B, C, D, E, F)$

- $f_1$: $ABC ightarrow D$
- $f_2$: $CE ightarrow B$
- $f_3$: $E ightarrow C$
- $f_4$: $BF ightarrow AC$
- $f_5$: $E ightarrow F$
- $f_6$: $F ightarrow E$