CSE 190D Spring 2017 Midterm Exam

Full Name:

Student ID:

Major:

INSTRUCTIONS

1. You have 50 minutes to complete this exam.

2. You can have up to one letter/A4-sized sheet of notes, formulae, etc. Apart from this, the exam is closed book/notes/electronics/peers.

3. Please wait until being told to start reading and working on the exam.

4. Please ensure that your writing is clear and legible!

<table>
<thead>
<tr>
<th></th>
<th>Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q 1</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Q 2</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Q 3</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Q 4</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>
Test:
1. When did the Pilgrims land at Plymouth Rock?

1620.

As you can see, I've memorized this utterly useless fact long enough to pass a test question. I now intend to forget it forever. You've taught me nothing except how to cynically manipulate the system. Congratulations.

They say the satisfaction of teaching makes up for the lousy pay.
Q 1. [12pts] Consider the following extendible hash index.

```
Bucket A
| 44* | 12* | 32* |
--- | --- | --- |
Bucket B
| 13* | 41* | 25* |
--- | --- | --- |
Bucket C
| 86* |
--- |
Bucket D
| 19* | 15* |
--- |
```

1. [6pts] Draw the index after the following sequence of update operations: delete 44*, insert 21*, delete 41*, and insert 34*.
2. [3pts] After performing the given sequence of updates, what is the minimum number of insert operations needed for the global depth to increase? Clearly circle the correct answer.

(a) 1       (b) 2       (c) 3       (d) 4       (e) 5       (f) 6

3. [3pts] Similarly, after performing the given sequence of updates, what is the minimum number of delete operations needed for the global depth to decrease instead? Clearly circle the correct answer.

(a) 1       (b) 2       (c) 3       (d) 4       (e) 5       (f) 6
Q 2. [12pts] For the following questions, clearly circle True or False.

1. In the delimiter-based record layout, we need to scan the record from the start even to retrieve a single field in the middle.

   True  False

2. In the unpacked page layout for fixed-length records, deleting records alters the record IDs of other records on the same page.

   True  False

3. The clock algorithm avoids the sequential flooding problem faced by LRU.

   True  False

4. Only one disk head can read/write at a given time even if the disk has many platters.

   True  False

5. The average rotational delay of a disk is a function of its RPM.

   True  False

6. It is possible to have an unclustered B+ tree index that follows the AltRecord alternative for data entries.

   True  False
Q 3. [14pts] Suppose we are sorting a relation with 100 million pages and we have 500 buffer pages for the external merge sort (EMS). A "pass" over the relation is defined as one read and write of the whole file. In all of the following, you have to include both the sort and merge phases. Clearly circle the correct answer for each of the following questions.

1. [2pts] How many passes will a naive 2-way EMS perform?
   (a) 22   (b) 24   (c) 26   (d) 28   (e) 30   (f) 32

2. [2pts] How many passes will a standard multi-way EMS perform, assuming we do not use any of the three improvements discussed in class?
   (a) 1   (b) 2   (c) 3   (d) 4   (e) 5   (f) 6

3. [2pts] How many passes will a multi-way EMS perform, assuming we use replacement sort for internal sorting?
   (a) 1   (b) 2   (c) 3   (d) 4   (e) 5   (f) 6

4. [3pts] How many passes will a multi-way EMS perform, assuming we use replacement sort for internal sorting along with blocked I/O with block sizes of 20 pages but no double buffering?
   (a) 1   (b) 2   (c) 3   (d) 4   (e) 5   (f) 6

5. [3pts] How many passes will a multi-way EMS perform, assuming we use replacement sort for internal sorting along with blocked I/O with block sizes of 20 pages and double buffering?
   (a) 1   (b) 2   (c) 3   (d) 4   (e) 5   (f) 6

6. [2pts] Which improvement will never cause the number of passes to increase?
   (a) Internal replacement sort   (b) Blocked I/O   (c) Double buffering
Q 4. [12pts] Where do you drink beer? Express the following queries in relational algebra, given the following database schema with three relations.

Likes (drinker, beer)
Frequents (drinker, bar)
Serves (bar, beer)

(Hint: You only need three operators: project, natural join, and set difference.)

1. [4pts] The **average** drinkers: which drinkers frequent *some* bar that serves *some* beer they like?

2. [4pts] The **prudent** drinkers: which drinkers frequent *only* bars that serve *some* beer they like?

3. [4pts] The **sad** drinkers: which drinkers frequent *only* bars that serve *no* beers they like?