Q 1. [12pts] Consider the following extendible hash index with 3 slots per bucket.

1. [6pts] Draw the index after the following sequence of update operations: delete 97*, insert 11*, delete 47*, and insert 34*.
2. [3pts] After performing the given sequence of updates, what is the minimum number of delete operations needed for any local depth to decrease? Clearly circle the correct answer.

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

ANSWER: (a) 1. Explanation: Delete 11* to trigger merge of buckets B and D.

3. [3pts] Similarly, after performing the given sequence of updates, what is the maximum number of insert operations possible before the global depth increases? Clearly circle the correct answer.

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

ANSWER: (e) 5. Explanation: Count the number of free slots: A has 1, B has 1, C has 1, and D has 2.

Q 2. [12pts] Suppose we are sorting a relation with 100 million pages and we have 1000 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 the naive 2-way EMS perform?

(a) 22  (b) 24  (c) 26  (d) 28  (e) 30  (f) 32

ANSWER: (d) 28

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) 2  (b) 3  (c) 4  (d) 5  (e) 6  (f) 7

ANSWER: (b) 3

3. [2pts] How many passes will a multi-way EMS perform, assuming we use replacement sort for internal sorting?

(a) 2  (b) 3  (c) 4  (d) 5  (e) 6  (f) 7
ANSWER: (b) 3

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 5 pages but no double buffering?

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

ANSWER: (d) 4

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 5 pages and double buffering?

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

ANSWER: (d) 4

Q 3. [14pts] For the following questions, clearly circle True or False.

1. Redistribution of index keys among siblings following an overflow at a node is never preferable for reorganizing a B+ tree index after an insert operation.
   
   FALSE

2. Using variable-length record layout for fixed-length records leads to more metadata overhead than using the packed record layout.
   
   TRUE

3. Upon a new page request, if all buffer frames have pinned pages, a typical buffer manager will pick a frame with the lowest pin count for replacement.
   
   FALSE

4. Random accesses typically achieve peak data transfer throughput from hard disks.
   
   FALSE

5. Primary indexes are always unique indexes.
   
   TRUE

6. The delimiter-based record format can store records longer than a page.
   
   TRUE
Q 4. [12pts] Do you fly? Given the following relational database schema, translate each given (SQL) query into an equivalent relational algebra query (or a sequence of queries) with the correct notation used for the relational operators. Drawing the relational algebra tree is also acceptable.

Airports (Code, City, State)
Flights (ID, FromCode, ToCode, Time)

Code is the primary key of Airports. Both FromCode and ToCode are foreign keys referring to Airports.Code.

1. [3pts] SELECT DISTINCT City, State FROM Airports, Flights WHERE ID = 123 AND FromCode = Code

\[ \pi_{\text{City, State}}(\sigma_{\text{ID}=123}(\text{Flights} \bowtie_{\text{FromCode}=\text{Code}} \text{Airports})) \]

2. [3pts] SELECT City, State, AVG (Time) FROM Airports, Flights WHERE Code = ToCode GROUP BY City, State

\[ \gamma_{\text{City, State, AVG(Time)}}(\text{Airports} \bowtie_{\text{Code}=\text{ToCode}} \text{Flights}) \]

3. [3pts] How many flights are there from San Diego (airport code SAN) to San Francisco (airport code SFO)?

\[ \gamma_{\text{COUNT(*)}}(\sigma_{\text{FromCode}=\text{"SAN"} \land \text{ToCode}=\text{"SFO"}}(\text{Flights})) \]

4. [3pts] How many airports can be reached from San Diego’s with exactly two flights?

\[ \gamma_{\text{COUNT(DISTINCT TC2)}}(\sigma_{\text{FromCode}=\text{"SAN"}}(\text{Flights} \bowtie_{\text{ToCode}=\text{FC2}} \rho_{\text{ID2,FC2,TC2,T2}}(\text{Flights}))) \]