Department of Computer Science
Core Graduate Comprehensive Exam
Spring 2020

- Answer the questions on the paper supplied.
- Answer question 1 or 2, Answer question 3 or 4 or 5. Answer question 6 or 7 or 8 or 9 or 10. Answer question 11 or 12 or 13 or 14. You should answer a total of four questions. Please Note: If you answer more than one question in one group, only the one with the LOWEST score will be counted.
- Start each question on a new page. Write on only one side of the paper.
- Write your SIX-DIGIT Texas State ID in the top right corner of each page of your answer. Do NOT put your name anywhere on the answers.
- Put the number of the question being answered in the top left corner of each answer page. Put the CORRECT question number to avoid missing your answer.
- If the answer to a question is written on more than one page, number the pages consecutively.

Group 1

1. CS 5329 Algorithm Design and Analysis
   
   { from Dr. Hwang }

   (a) Use iteration method or recursion tree method to determine the total running time \( T(n) \) of the recurrence (formula) :
       \[
       T(n) = 3 \cdot T\left(\frac{n}{4}\right) + n
       \]

   (b) Use iteration method or recursion tree method to determine the total running time \( T(n) \) of the recurrence (formula) :
       \[
       T(n) = 2 \cdot T\left(\frac{n}{2}\right) + n
       \]

2. CS 5329 Algorithm Design and Analysis
   
   { from Dr. Metsis }

   Binary Heaps
   
   A Min Heap is a heap that obeys the min-heap property. That is, assuming an array \( A \), for each element at position \( i \) of the array: \( A[\text{Parent}(i)] \leq A[i] \), where \( \text{Parent}(i) \) is a function which returns the position of the parent of element \( i \) in the heap tree.

   Assuming the binary Min Heap below, answer the following questions:

   ![Binary Heap Diagram]

   a) Show the resulting heap tree and array contents after inserting the element -2 to the heap.

   b) Show the resulting heap tree and array contents after deleting the minimum element of the original heap (before inserting -2).

   c) Using pseudocode, write a subroutine that prints out the elements of the heap in a postorder manner, i.e. for each node in the heap tree, the values of its children are printed first, and then the value of the node is printed. Start from the first element of the heap array. You can assume that the heap has a given size \( n \).
3. **CS 5346 Advanced Artificial Intelligence**  
   *from Dr. Ali*

   (a) A database of a county bank’s loan department is represented as positive and negative instances in the following table:

<table>
<thead>
<tr>
<th>Individual</th>
<th>INC</th>
<th>BAL</th>
<th>APP</th>
<th>RATING</th>
<th>OK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

   Derive a set of rules that covers all and only the positive instances.

   (b) (i) An algorithm has the following evaluation function:

   \[ f(n) = (2-w)g(n) + w \cdot h(n) \]

   For what kind of search does this perform for \( w = 0, w = 1, w = 2 \)? For what values of \( w \) is it complete and optimal?

   (ii) Another algorithm uses the following evaluation function:

   \[ f(n) = (2-W)g(n) + W/(2-W) \cdot h(n) \]

   Assuming \( h(n) \) is admissible, what value of \( W \) will make it behave exactly like \( A^* \)?

4. **CS 5391 Survey of Software Engineering**  
   *from Dr. Chen*

   Explain the five levels in the Capability Maturity Model. Please skip if you do not remember what they are.

5. **CS 5391 Survey of Software Engineering**  
   *from Dr. Palacios*

   You have been hired to lead a new software quality assurance (SQA) group in a medium-sized software development company that specializes in a Web app for online voting and elections. Currently, the main quality concern about the company’s software product is the security protections of its electoral data.

   As you know, SQA includes several areas of concern and activities that focus on the management of software quality. Indicate three SQA areas and/or activities that you would focus on right away, as you set up your working group, and also indicate how you would go about implementing them in the software development group.

Group 3

6. **CS 5306 Advanced Operating Systems**  
   *from Dr. Tamir*

   Assume a single core system implementing an intra-core preemptive HRRN scheduling policy with a slice size of 1 second. Further assume that at time \( T \) there are 3 tasks \( \{T_0, T_1, T_2\} \) in the Ready Queue of the core with no task in the execution slot of the core. Additionally, assume that the tasks are compute-bound with no I/O whatsoever. Let \( \{P_0, P_1, P_2\} \) be the remaining execution time of \( \{T_0, T_1, T_2\} \) respectively and let \( \{P_3, P_4, P_5\} \) be the current wait time of \( \{T_0, T_1, T_2\} \) respectively, where:

   - \( P_0 = 1 \)
   - \( P_1 = 4 \)
   - \( P_2 = 2 \)
   - \( P_3 = 1 \)
   - \( P_4 = 2 \)
   - \( P_5 = 3 \)

   1) Clearly describe the state of the system in each of the first 20 seconds following time \( T \).
7. **CS 5306 Advanced Operating Systems**
   
   Design a distributed algorithm to synchronize physical clocks in a distributed system in a room.

8. **CS 5332 Data Base Theory and Design**

   (a) Specify all functional dependencies for the following library problem

   (b) Create an E/R model to model the library problem.
   
   (Must include pk’s, fk’s and relationships and all tables must be normalized.)

   Problem: The library provides books to borrowers. Each book is described by title, edition, and year of publication, and is uniquely identified using the ISBN. Each borrower is described by his or her name and address and is uniquely identified using a borrower number. The library provides one or more copies of each book and each copy is uniquely identified using a copy number, status indicating if the book is available for loan, and the allowable loan period for a given copy. A borrower may loan one or many books, and the date each book is loaned out and is returned is recorded. Loan number uniquely identifies each book loan.

9. **CS 5332 Data Base Theory and Design**

   The customer database consists of the relations defined in the four tables shown below:

   ```
   Customer(custNo, custName, state, phone)
   Item(itemNo, itemName, itemPrice, qtyOnHand)
   Invoice(invNo, invDate, custNo)
   Invitem(invNo, itemNo, qty)
   ```

   Primary keys are underlined. Foreign keys are shown in italic font. Primary keys that are also foreign keys are shown in underline italic font. For example, the invNo and itemNo attributes in Invitem relation is the primary key for Invitem and is also foreign keys of tables Invoice and Item. Below are sample data from each of the above relation.

   **Customer table**
   
<table>
<thead>
<tr>
<th>custNo</th>
<th>custName</th>
<th>state</th>
<th>phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>Garcia</td>
<td>TX</td>
<td>732-555-1000</td>
</tr>
<tr>
<td>212</td>
<td>Parkin</td>
<td>NY</td>
<td>212-555-2000</td>
</tr>
<tr>
<td>225</td>
<td>Ellen</td>
<td>NJ</td>
<td>973-555-3333</td>
</tr>
<tr>
<td>239</td>
<td>Bayer</td>
<td>FL</td>
<td>401-555-7777</td>
</tr>
</tbody>
</table>

   **Item table**
   
<table>
<thead>
<tr>
<th>itemNo</th>
<th>itemName</th>
<th>itemPrice</th>
<th>qtyOnHand</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Screw</td>
<td>5.00</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>Nut</td>
<td>2.25</td>
<td>110</td>
</tr>
<tr>
<td>3</td>
<td>Bolt</td>
<td>3.99</td>
<td>75</td>
</tr>
<tr>
<td>4</td>
<td>Hammer</td>
<td>10.00</td>
<td>125</td>
</tr>
<tr>
<td>5</td>
<td>Washer</td>
<td>6.00</td>
<td>300</td>
</tr>
<tr>
<td>6</td>
<td>Nail</td>
<td>0.99</td>
<td>400</td>
</tr>
</tbody>
</table>

   **Invoice table**
   
<table>
<thead>
<tr>
<th>invNo</th>
<th>invDate</th>
<th>custNo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001</td>
<td>05-Sep-07</td>
<td>212</td>
</tr>
<tr>
<td>1002</td>
<td>17-Sep-07</td>
<td>225</td>
</tr>
<tr>
<td>1003</td>
<td>17-Sep-07</td>
<td>239</td>
</tr>
<tr>
<td>1004</td>
<td>18-Sep-07</td>
<td>211</td>
</tr>
<tr>
<td>1005</td>
<td>21-Sep-07</td>
<td>212</td>
</tr>
</tbody>
</table>
Write SQL to answer the following queries:

(a) Find items which have the itemPrice between $5 and $10?
(b) Count the number of items ordered in each invoice.
(c) Find customers who are not from "NY" or "NJ"
(d) Find items that are cheaper than nut. You must display the itemName and the associated itemPrice. Beware that nut can be written in both upper and lower case.
(e) Find the invoices for customers in which three or more items are ordered.
(f) List the invoices, customer names and the items purchased by them. The result should display invoices No, customer names and item names together ordered by the customer name.

10. **CS 5310 Network and Communication Systems**
   { from Dr. Peng }
   (a) Describe the steps and protocols used when you send an email message using a computer in CS lab to a friend of yours who uses an email server on a campus in California.
   (b) Describe briefly why network layer is not an end-to-end layer.

Group 4

11. **CS 5338 Formal Languages**
   { from Dr. Gao }
   (a) Draw a PDA $M$ that recognizes language $L(M) = \{b^n a^n : n \geq 0\}$.
   (b) Use operators $= $ (meaning equal) and $<$ (meaning less than) to rank the following abstract machines in terms of computability.
   
   DTM (deterministic Turing Machine)
   NDTM (nondeterministic Turing Machine)
   DFSM (deterministic FSM)
   NDFS (nondeterministic FSM)
   DPDA (deterministic PDA)
   NDPDA (nondeterministic PDA)
   
   (c) From the following abstract machines, underline the ones that must halt at all times.
   
   DFSM
   NDFS
   NDFS without $\epsilon$-transitions
   DPDA
   NDPDA
   NDPDA without $\epsilon$-transitions
   DTM
   NDTM
   
   (d) The P versus NP problem is a major unsolved problem in computer science. Assuming $P \neq NP$, use a Venn diagram to show the logical relations for sets P, NP, and NP-complete.
12. **CS 5338 Formal Languages**  
*from Dr. Singh*

What is Pumping Lemma for regular languages? Using Pumping Lemma prove \( L = \{0^n1^n | n \geq 1 \} \) is not regular.

13. **CS 5318 Design of Programming Languages**  
*from Dr. Shi*

(a) Consider the simple grammar for arithmetic expressions:

\[
E \rightarrow T \mid E + T \\
T \rightarrow P \mid P * T \\
P \rightarrow i \mid (E)
\]

i. What is the associativity of each operator?
ii. Give a leftmost derivation of expression: \( i + i * i + i \) and its parse tree.
iii. Re-write the grammar using EBNF (extended BNF).
iv. Write a recursive-descent parser based on the grammar written in EBNF.

14. **CS 5351 Parallel Processing**  
*from Dr. Burtscher*

Assume we want to parallelize the outer loop of the following code on a shared-memory system without caches.

```java
for (int i = 0; i < n; i++) {
    int a = i;
    for (int j = 0; j < i; j++) {
        // do O(1) work here that reads variable a
        a += 3;
        // do some other O(1) work here that reads variable a
    }
}
```

1) Which of the following three workload partitioning schemes results in the least amount of load imbalance and why: blocked, cyclic, or block-cyclic?

2) If the system had a data cache, would a different workload partitioning scheme be better? If not, why not? If so, under what conditions and why?

3) Assuming that \( n \) is a large positive number and that each \( i \)-loop iteration is assigned to a separate thread (and no other parallelization is used), what is the parallel efficiency?

4) Assume each thread has a private variable called “rank” that contains the thread’s unique rank. Assume further that there is a shared variable called “threads” that contains the number of running threads and that all ranks are in the range 0 through \( \text{threads} - 1 \). Using these variables, rewrite the above code such that it assigns the \( i \)-loop iterations cyclically to the threads.

5) Rewrite the above code such that the loop-carried data dependence on variable “\( a \)” is eliminated but the code maintains its semantics.