

Preview Test: INFS7901 Semester One End-of-semester Examination 2022

Test Information

Description Undertaking this online examination deems your commitment to UQ's academic integrity pledge as summarised in the following declaration:

"I certify that I have completed this examination in an honest, fair and trustworthy manner, that my submitted answers are entirely my own work, and that I have neither given nor received any unauthorised assistance on this examination".

Instructions You need to answer all of the questions in the Blackboard Test.

If you believe there is missing or incorrect information impacting your ability to answer any particular question, please state this when answering that question.

You must do all your working in the Blackboard test.

Timed Test This test has a time limit of 2 hours and 10 minutes. This test will save and be submitted automatically when the time expires.

Warnings appear when **half the time, 5 minutes, 1 minute**, and **30 seconds** remain.
[The timer does not appear when previewing this test]

Multiple Attempts Not allowed. This test can only be taken once.

Force Completion This test can be saved and resumed at any point until the time has expired. The timer will continue to run if you leave the test.

Your answers are saved automatically.

QUESTION 1

0 points

Save Answer

You may access the below lecture materials when completing your exam. Please note there are no lecture materials from Week 9.

[1 ER Diagram.pdf](#)

[2 Relational Model.pdf](#)

[3 FD and Normal Form.pdf](#)

[4 SQL\(1\).pdf](#)

[5 SQL\(2\).pdf](#)

[6 Asymptotic Analysis.pdf](#)

[7 Sorting.pdf](#)

[8 Trees.pdf](#)

[10 Hashing.pdf](#)

[11 Indexing.pdf](#)

[12 Query Optimization.pdf](#)

Agree

Disagree

QUESTION 2

3 points

Save Answer

In a schema with attributes A, B, C, D, and E, the following set of functional dependencies are given:

A→B

A→C

CD→E

B→D

E→A

Which of the following functional dependencies is not implied by the above set? Choose one answer only.

A. CD→AC ✓

B

B. BD→CD

C. AC→BC

D. BC→CD

QUESTION 3**3 points**

Save Answer

Which statement is **true** about SQL? Choose one answer only.

- A. Both WHERE and HAVING clauses should not include aggregates. *where not include aggregates*
- B. Conditions "IN" and "= ANY" shouldn't be used interchangeably *Some time can be interchangeably at any time due to inequivalence*
- C. The "WHERE r1.age > r2.age" condition will return True by default if r2.age = NULL. *in null or is not null*
- D. Two tables that do not share the same attribute names can still be union-compatible. *Some domain is shared*

QUESTION 4**3 points**

Save Answer

Which sorting algorithm is being implemented by the following "secretFunction"? Choose one answer only.

```
def secretFunction(aList):
    for i in range(1, len(aList)):
        tmp = aList[i]
        k = i
        while k > 0 and tmp < aList[k - 1]:
            aList[k] = aList[k - 1]
            k = k - 1
        aList[k] = tmp
```

D

- A. Quick Sort *PIVOT*
- B. Selection Sort
- C. Merge Sort
- D. Insertion Sort

QUESTION 5**3 points**

Save Answer

Which of the following time complexities has the fastest growth rate?
Choose one answer only.

- A. $O(n^2)$
- B. $O(n \log n)$
- C. $O(2^n)$
- D. $O(\log n)$

C

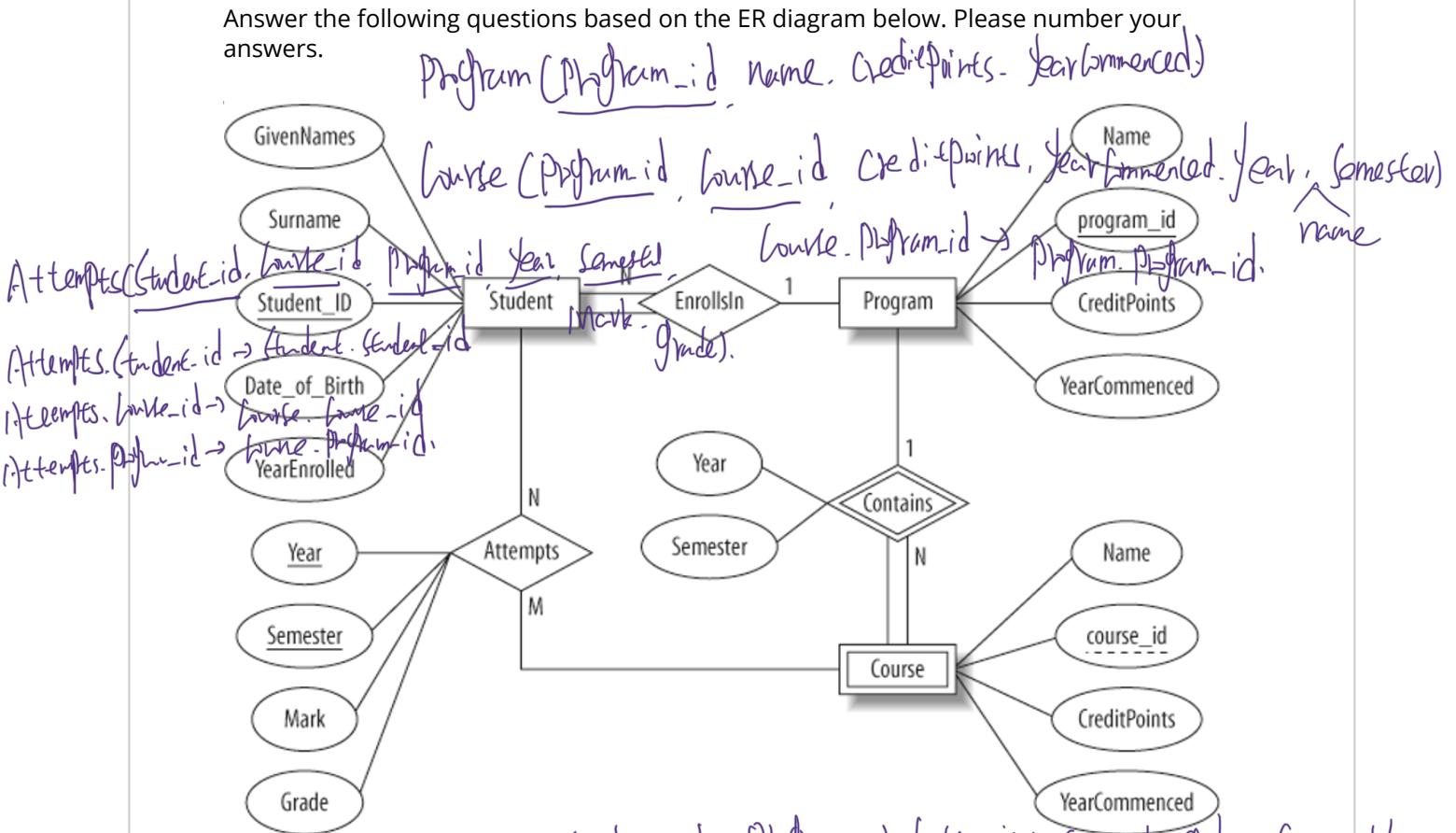
$n! > 2^n > n^3 > n \log n > n > \log n$

QUESTION 6

10 points

Save Answer

Answer the following questions based on the ER diagram below. Please number your answers.



NO, Student to Program relation is 1. So each student can only enrol one program.

- (2 marks) Can a student enrol in multiple programs at the same time (Yes/No)? Explain in one sentence.

- (2 marks) Can we have programs in the Program table that are not enrolled by any student (Yes/No)? Explain in one sentence.

- (2 mark) Give the minimal key for the Course entity. program_id, course_id

- (4 marks) Map this ER diagram into a relation schema. Make sure to underline attributes of the primary key, and write down all foreign keys in the format of "Relation_X.Attribute_A → Relation_Y.Attribute_B". Marking foreign keys in bold font is not mandatory.

* The arrow sign can be copied from here: → Student (Student_id, YearEnrolled, Date_of_Birth, Surname, GivenName, Program_id)

For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).

B	I	U	S	Paragraph	Open Sans, sa...
10pt	12pt	14pt	16pt	A	I
X	□	□	Q	↶ ↷	≡ ≡ ≡ ≡ ≡ ≡
🔗	✖	✖	✖	– ±	ABC ...
P	0 WORDS POWERED BY TINY				

QUESTION 7

10 points

[Save Answer](#)

Answer the questions based on the relation instances given in the table below. Please number your answers.

A	B	C	D
1	John	IT	1
2	Mark	Business	1
3	Mary	Health	1
4	Jordan	IT	1
5	Mary	Art	0
6	Jordan	Law	0

1. (2 marks) Check the given instances, and specify whether each of the following dependencies is possible or not:

$A \rightarrow C$ (possible/impossible)
 $A \rightarrow D$ (possible/impossible)
 $C \rightarrow B$ (possible/impossible)
 $B \rightarrow D$ (possible/impossible)

{A, C}

2. (2 marks) Assuming that two functional dependencies $A \rightarrow B$ and $C \rightarrow D$ hold in the given schema, find the minimal key for the relation.

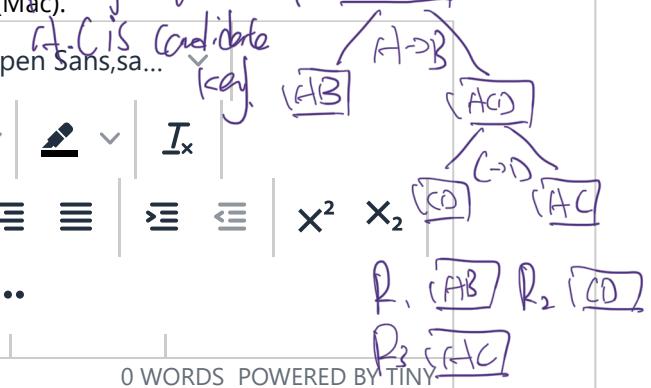
$$\begin{aligned} A^+ &= \{A, B, D, C\} \\ B^+ &= \{B\}. \end{aligned}$$

3. (2 marks) Assuming that three functional dependencies $A \rightarrow B$, $D \rightarrow C$ and $AB \rightarrow D$ hold in the given schema, what is the closure of A? What is the closure of B?

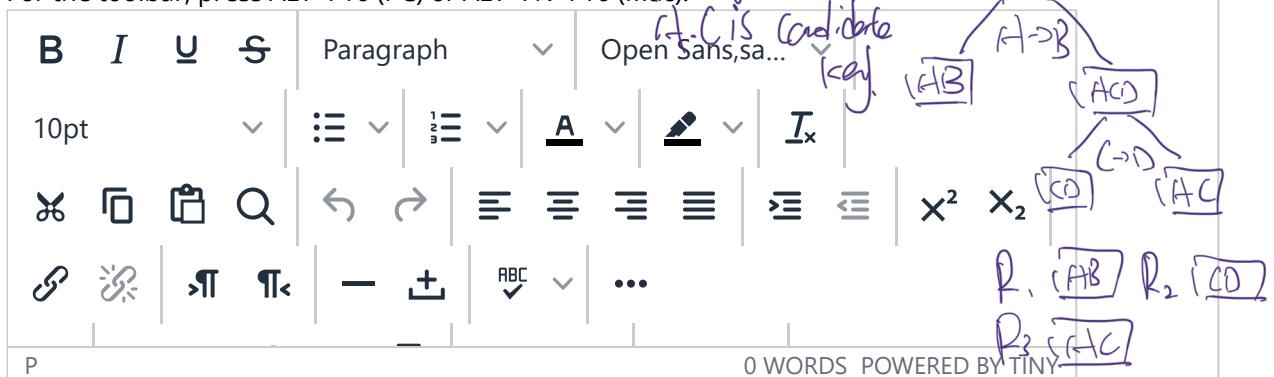
$$f(A.C)^H = \{A, B, C, D\}$$

ABC

4. (4 marks) Assuming only two functional dependencies $A \rightarrow B$ and $C \rightarrow D$ hold in the given schema. Is this relation in BCNF? If yes, give your reason. If no, decompose it to meet BCNF (only provide all the resulting tables).



For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).



QUESTION 8

10 points

Save Answer

Consider the given relational schema, in which each patient is treated by a doctor and has an emergency contact. A patient is either a normal patient or emergency patient.

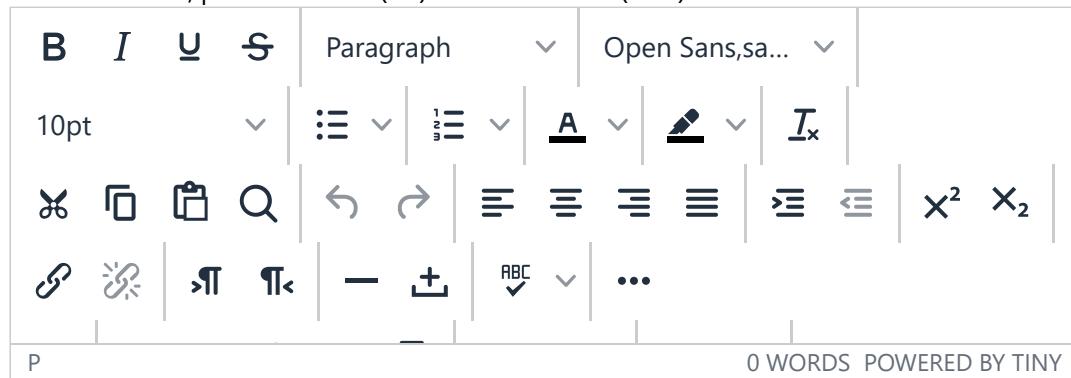
Tables:
Patient (<u>patientID</u> , name, DoB, doctorID)
Doctor (<u>doctorID</u> , name)
EmergencyContact (<u>patientID</u> , contact name, phone number)
NormalPatient (<u>patientID</u>)
EmergencyPatient (<u>patientID</u> , symptom)
Foreign Keys:
Patient.doctorID → Doctor.doctorID
EmergencyContact.patientID → Patient.patientID
NormalPatient.patientID → Patient.patientID
EmergencyPatient.patientID → Patient.patientID

- ① Select Symptom. Count(PatientID). from EmergencyPatient group by Symptom;
 - ② Select Symptom from EmergencyPatient EP join Patient P on EP.PatientID = P.PatientID
where P.DOB > '1970-01-01' ;
 - ③ Select Doctor.name from Doctor D join Patient P on D.doctorID = P.doctorID join NormalPatient NP on NP.patientID join EmergencyPatient EP on EP.patientID = P.patientID = NP.patientID
where EP.patientID is Null and NP.patientID is not Null;
 - ④ Select name from doctor D join PatientP on D.doctorID = P.patientID
join EmergencyEP on EP.patientID = P.patientID
where Symptom = 'Faint' having Count(EP.patientID) > 5;

Based on the schema, write the following queries in SQL. Please number your answers.

1. (2 marks) Count the number of emergency patients for every symptom.
 2. (2 marks) Find the symptoms of emergency patients who were born (i.e., DoB attribute) after 1 Jan 1970 (1970-01-01 in date datatype).
 3. (2 marks) Find the names of doctors who have treated at least one normal patient but no emergency patients.
 4. (2 marks) Find the names of doctors who have treated at least five emergency patients that have the symptom 'faint' (string datatype).
 5. (2 marks) For each symptom, find the most experienced doctor who has given the most amount of treatments.

For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).



QUESTION 9**10 points**

Save Answer

Answer the following questions about asymptotic analysis. Please number your answers.

1. (4 marks) For each of the following pairs of runtimes, determine which one is asymptotically **faster** (i.e., has lower computational complexity), or whether the two complexities are similar. Note that all log functions have base 2. Simply write down $T(A)$ or $T(B)$ for each pair.

Pair 1: $T(A) = \frac{n}{\log n}$ vs. $T(B) = \log n$ $T(B)$

Pair 2: $T(A) = 2^{\log n} + n^2$ vs. $T(B) = \frac{2^n}{n^2}$ $T(A)$

Pair 3: $T(A) = n^2 + n!$ vs. $T(B) = \log n \cdot n!$ $T(A)$

Pair 4: $T(A) = n^2 + 16 \log n$ vs. $T(B) = 2n^2 + n^3$ $T(A)$

2. (6 marks) Determine the running time of the two following code snippets using big-O notation. Briefly justify your answer.

*If you would like to save time in formatting your answers, feel free to describe them in natural language, e.g., $O(\text{square root of } n)$.

Code 1:

```
int i, j, k = 0;
for (i = n / 2; i <= n; i++) {
    for (j = 2; j <= n; j = j * 2) {
        k = k + n / 2;
    }
}
```

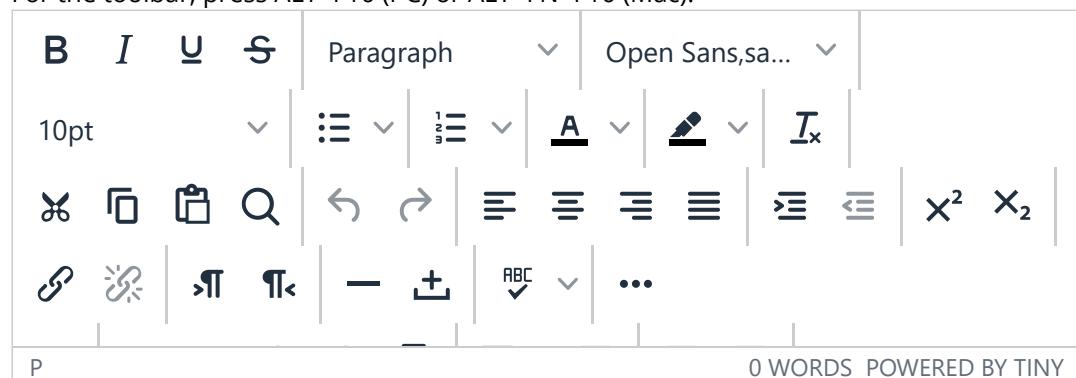
$\Theta(n \log n)$

Code 2:

```
int a = 0;
for (i = 0; i < N; i++) {
    for (j = N; j > i; j--) {
        a = a + i + j;
    }
}
```

$\Theta(n^2)$

For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).



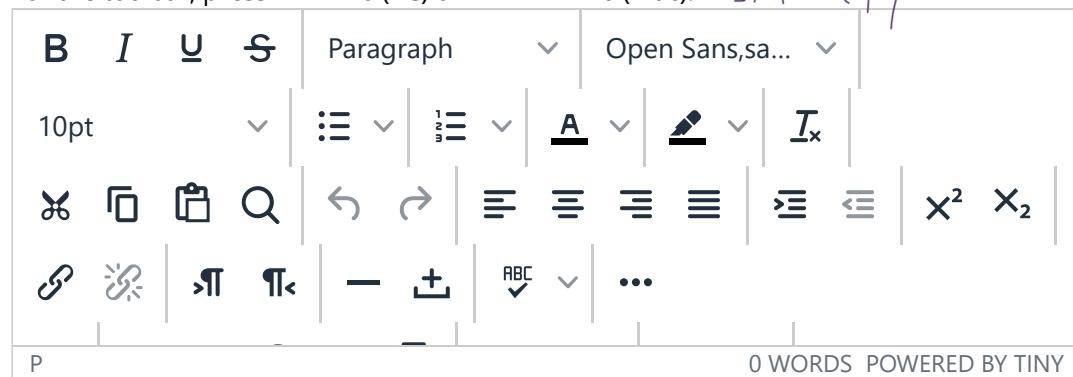
QUESTION 10**10 points**

Save Answer

Answer the following questions about sorting. Please number your answers.

1. (2 marks) Explain what makes a sorting algorithm stable. *after ordering the position for values are not changed*
2. (2 marks) When trying to sort an integer array with a large number of elements (e.g., billion-scale), which sorting algorithm would you choose from Selection/Insertion/Merge/Quick Sort to ensure the lowest time cost in *worst cases*? Which would you choose to ensure the lowest time cost in **best cases**? Justify both answers using asymptotic complexity. *insertion sort O(n)*
3. (2 marks) Suppose the long array mentioned above has been sorted. However, a small number of new elements (e.g., 5 elements) have been appended to the end of this array. To keep all elements sorted in this updated array, what sorting algorithm would you choose from the four options provided above? Provide an explanation on your answer. *insertion sort O(n)*
4. (2 marks) Sort array [6, 3, 15, 7, 2, 9] using selection sort. Only write down the updated array every time there is a swap operation. *2 3 15 7 6 9
2 2 6 7 15 9*
5. (2 marks) After the first iteration of quick sort the array is partitioned as [3, 4, 5, 13, 11, 19]. List all number(s) that could have been chosen as the pivot for the first iteration. *2, 4, 5, 19*

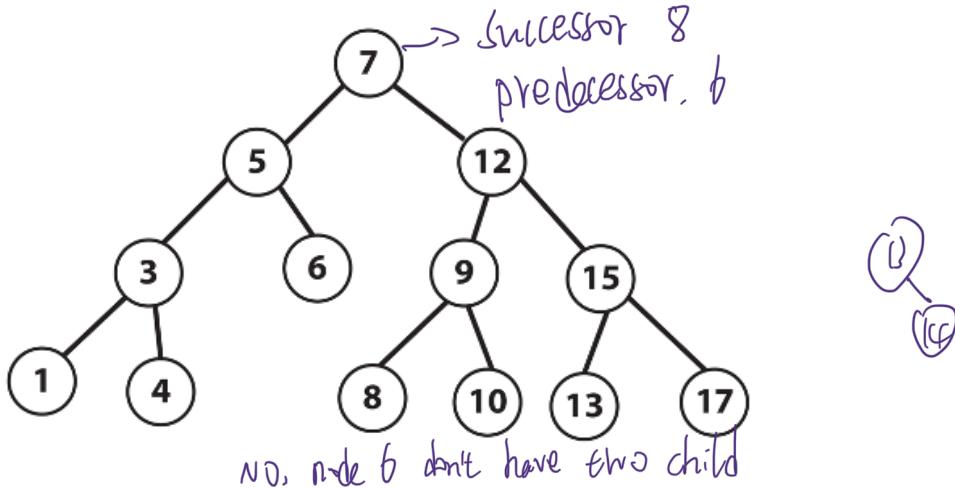
For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).



QUESTION 11**14 points**

Save Answer

Answer the following questions regarding the given Binary Search Tree. Please number your answers.



1. (3 marks) Is this tree complete? Is this tree almost/nearly complete? Provide justifications on both of your answers. *No, the last level is not left to right.*

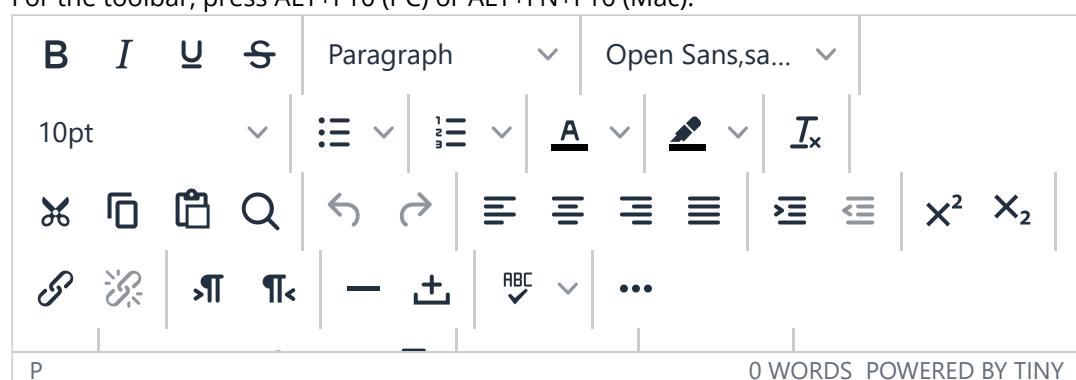
2. (2 marks) Give the successor and the predecessor of the root node.

3. (3 marks) List the nodes that will be visited by order if we are searching for 10. *7 → 1 → 9 → 10*

4. (3 marks) If we are adding 14 to the tree, what changes (if any) will be made to this tree? Alternatively, if we are adding 3 to the tree, what changes (if any) will be made to this tree? Briefly describe. *Do nothing.*

5. (3 marks) If we are deleting node 15, what changes (if any) will be made to this tree? Briefly describe. *use 13 to replace 15*

For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).



QUESTION 12**12 points**

Save Answer

Answer the following questions about indexing. Please number your answers.

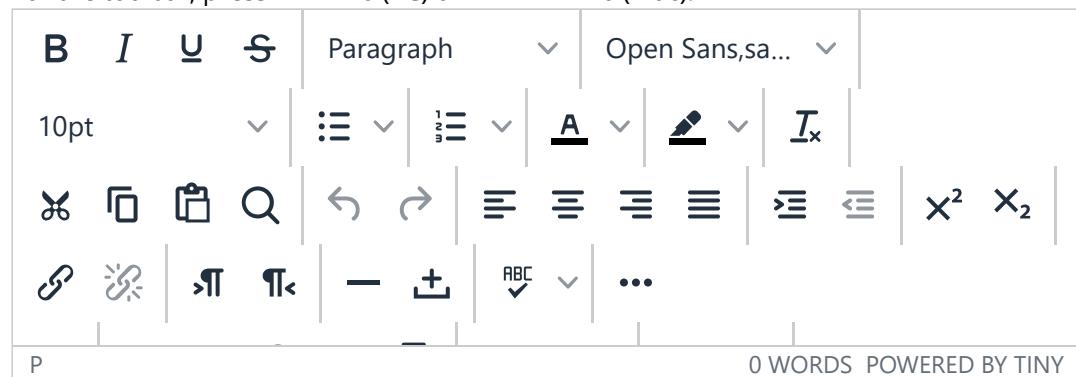
Given the schema below, what types of indexes will you design for each of the four frequent queries below? Two choices need to be made for each of the questions. Firstly, choose from primary, secondary, or clustering indexes. Secondly, choose from B+ Tree, hash-based, or grid file indexes.

Tables:
Student (<u>sid</u> , sname)
Course (<u>cid</u> , cname, level)
Certificate (<u>sid</u> , <u>cid</u> , date)
Foreign Keys: Certificate.sid → Student.sid Certificate.cid → Course.cid

1. Secondary, hash-based.
2. Primary, grid-file
3. Secondary, B+ Tree
4. Primary, hash-based.

1. (3 marks) select * from Student where sname = "A Name"
2. (3 marks) select * from Certificate where cid = "A Course ID" and pid = "A Student ID"
3. (3 marks) select * from Course where level > 2
4. (3 marks) select * from Course where cid = "A Course ID"

For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).



QUESTION 13

12 points

Save Answer

Consider the following relational schema, and answer questions about query optimization. Please number your answers.

Tables:

Customer(cid, cname, age)

Product(pid, pname)

Purchase(cid, pid, date, price)

Foreign Keys:

Purchase.cid → Customer.cid

Purchase.pid → Product.pid

* Feel free to copy and paste symbols from here: $\sigma \pi \cup \cap - \times \rho \bowtie \leftarrow \rightarrow /$

* To add subscripts, click the "... " button on the upper right corner of your textbox, and use the "X₂" button.

1. (6 marks) Write the following queries in relational algebra.

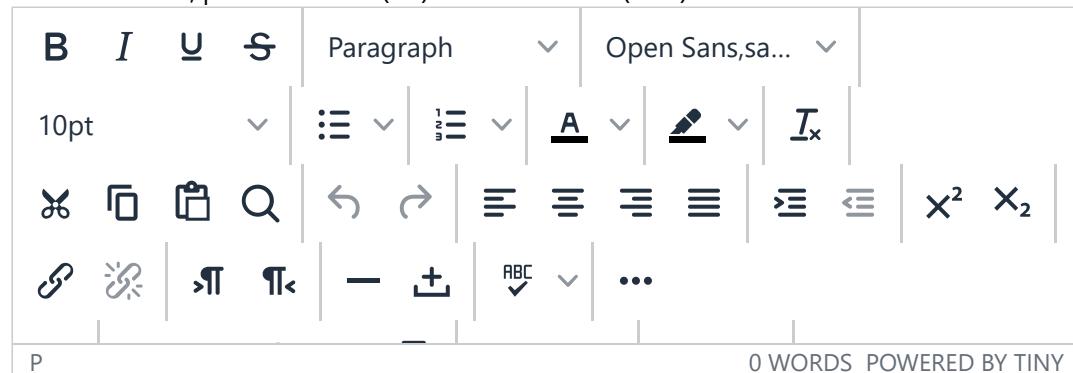
1a. Find the name of all products whose pid is smaller than 12345.

1b. Find the name of products that have been purchased.

1c. Find the name and age of all customers that purchased a product with a price greater than 900 (integer type) between 2022/01/01 and 2022/06/13.

2. (6 marks) Write the relational algebra of the **optimised** query plan for the following query: find the name and age of all customers that purchased a product with pname = "phone" on date < 2022/06/13. Be clear about the order of each relational operation. You are not required to draw the query tree, and you are not required to specify the algorithms used for each relational operation.

For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).



QUESTION 14

0 points

Save Answer

Please use this space if you believe there is missing or incorrect information that impacted your ability to answer any question. Please state which questions this is related to.

For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).

