（装订线内不要答题）

复旦大学软件学院

2022-2023 学年第一学期期末考试试卷

☐ A 卷 ☑ B 卷 ☐ C 卷

课程名称：数据库设计 课程代码：SOFT130015

开课院系：软件学院 考试形式：开卷

姓名：＿＿ 学号：＿＿ 专业：＿＿

提示：请同学们秉持诚实守信宗旨，谨守考试纪律，摒弃考试作弊。学生如有违反学校考试纪律的行为，学校将按《复旦大学学生纪律处分条例》规定予以严肃处理。

题号	1	2	3	4	5	6	7	8	总分
得分									

Answer all questions on the Answer Sheets.

Questions marked with "**[H]**" are only for the students enrolled in the Honor Class.

Time allowed: 2 hours

I.(9 Points) Explain the following concepts in your own words.

(i)(3 Points) Candidate Key

(ii)(3 Points) Multilevel Index

(iii)(3 Points) The Isolation Property (of a transaction)

II.(18 Points) Answer the following questions.

(i)(6 Points) Please show with an example where the third normal form (3NF) is purposefully violated.

(ii)(6 Points) Please explain how a bitmap index works to improve query performance.

(iii)(6 Points) ①Please explain what will happen if two transactions T_1 and T_2 are scheduled as follows. ②How should the schedule be modified so that the execution of the two transactions meet the consistency requirement.

T_1	T_2
read (A)	
write (A)	
	read (A)
	commit
abort	

III.(10+4 Points) Query writing.

(i)(3+4 Points) Consider the following schemas: R= (A, B, D) and S= (B, C, E). $r(R)$ and $s(S)$ are relations.

Please write:

(a) a statement in SQL,

(b) **[H]** an expression in tuple relational calculus, and

(c) **[H]** an expression in domain relational calculus,

so that all of them are equivalent to the following query:

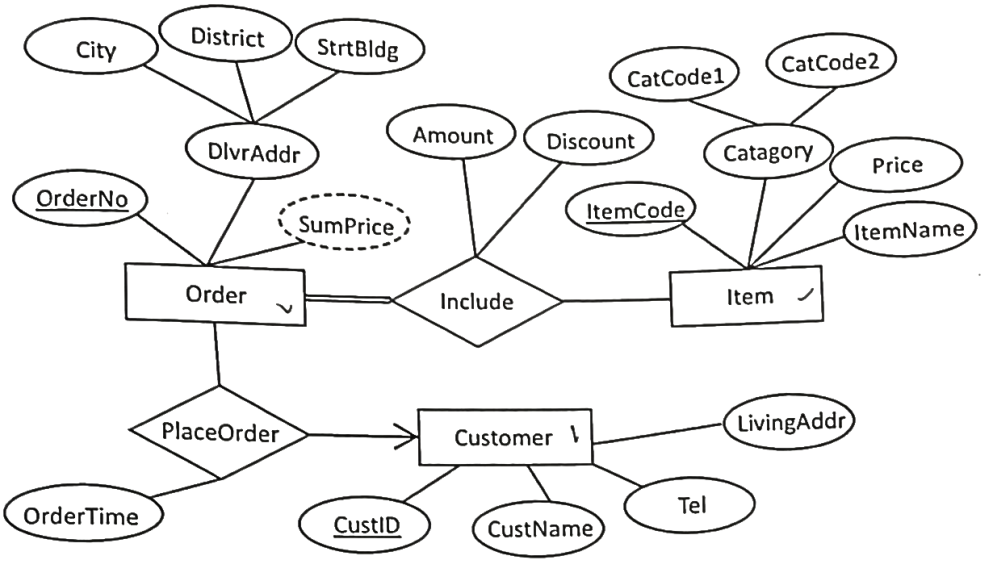
$$\prod_{A,B,D,E}(\sigma_{C='A'\wedge D>1000}(r \bowtie s))$$

(ii)(3 Points) Please write an expression in relational algebra to find the average value of D for each value group of C with the join of r and s on attribute B.

(iii)(4 Points) Given schemas R=($\underline{A, B}$, C) and S=($\underline{B, C, E}$), where the primary keys are underlined. $r(R)$ and $s(S)$ are relations. Please write a query in SQL, which is equivalent to the following expression.

$$r \div \prod_{B,C}(\sigma_{E=0}(s))$$

IV. (22 Points) Answer all questions based on the given E-R diagram.



(i)(12 Points) Try to convert it into ***minimal*** relation schemas. Each schema should have the form like *relation1(attribute1, attribute2, attribute3)* where the primary key is underlined.
(ii)(6 Points) Please write **a query in SQL** to calculate the correct value of *SumPrice* of a specific *Order* (for example, 'A12345') based on the order details in the *Items*. Display the order number and the corresponding SumPrice in the result.
(iii)(4 Points) If the *SumPrice* is maintained by a trigger, please tell the design of the trigger and then discuss the **benefits and/or drawbacks**. You do not need to write the implementation details of the trigger but only to describe the functions and triggering condition(s) of it.

V.(22 Points)

Given the schema $R = (A, B, C, D, E)$ and the following set F of functional dependencies for R.
$A \rightarrow BC$
$B \rightarrow DE$
$CD \rightarrow A$
$AD \rightarrow E$

(i)(2 Points) Is AC a candidate key of R? How do you know that?

(ii)(3 Points) Is R in 3NF? Why?

(iii)(4 Points) Please compute a canonical cover F_c of F and show your steps.

(iv)(4 Points) Please give a lossless decomposition into 3NF of schema R, preserving functional dependencies.

(v)(4 Points) Please give a lossless decomposition into BCNF of schema R, and then show whether functional dependencies are preserved in the decomposition.

(vi)(5 Points) If we decompose *R* into *R1*(*A, C*), *R2*(*A, D, E*), *R3*(*B, D, E*), please show whether the decomposition is lossless.

VI.(19 Points) Consider the following description of a device management system and answer the questions that followed.

The device management system is used to manage several types of electronic devices, such as server computers, desktop computers, laptop computers, displays, portable projectors, portable hard disks, printers, etc. Devices are identified by registered IDs. The system also records other information about the device, such as name, type, date of bought, etc.

The system records the relationship between devices and employees. One device is related to at most one employee at the same time. If a device is not related to any employee, we say the device is available for employees to use. If a device is for sharing (e.g., a shared printer), the related employee should be the device manager, a specific employee who manages information of all devices.

When a device needs repairing, the related employee shall fill out an electronic form of repair request which records the device ID and the reason of repairing. The device manager deals with such requests and do the repair work if he is capable to do the repair. Otherwise, he contacts the device supplier for warranty. The request will not close until the repair is done or denied.

Please answer the following questions based on the above descriptions.

(i)(3 Points) What do you think could be the most frequently used queries (data requirements) in terms of everyday device management? Please list at least three of them in natural language. [Hint: such as "list the open repair requests" and you may imagine what are valuable for the device manager or other employees.]

(ii)(5 Points) Please perform a conceptual design by constructing an E-R diagram based on the above description. Please include all attributes that are mentioned in the description. Add any attribute if you think you have to.

(iii)(3 Points) Please derive appropriate relation schemas based on ERD designed in (i). Each schema should be in the form *relation*(*attribute*, *attribute2*, *attribute3*), where primary key is underlined. Should you make any assumptions, document them as well.

(iv)(8 Points) Please show how your design will work for a large company (e.g., with 50,000+ employees in tens of branches around the world) where the history of employee-device relationship should be tracked. List potential problems and show how your design addresses them or how you should modify your design to address them. [Hint: The "problems" may be either on a database design level or on an application design level, or both.]

VII. [H] (+6 Points) Do you have any comments or suggestions on designing a database system to store the detection results of Automatic Static Analysis Tools (SATs)? Please explain your ideas briefly.