

**Đề Kiểm tra Cuối kỳ – Học kỳ 2 – Năm học 2017-2018**  
**HỆ CƠ SỞ DỮ LIỆU (MSMH: CO2013)**  
 Ngày thi: 05/06/2018 - Thời gian: 90 phút  
**Đề số: 0001**

Cán bộ ra đề:

Chủ nhiệm bộ môn duyệt đề:  
 Ngày 1 tháng 6 năm 2018

Họ tên sinh viên: \_\_\_\_\_ MSSV: \_\_\_\_\_

Lưu ý: - Sinh viên **ĐƯỢC** được sử dụng tài liệu dạng **GIẤY**.

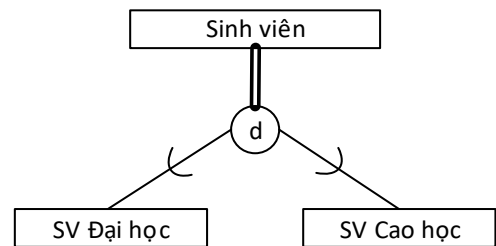
- Sinh viên **làm phần trắc nghiệm trên phiếu trắc nghiệm** và **phần tự luận trực tiếp trên đề thi**.
- Giám thị không cần giải thích gì thêm.
- Đề thi gồm 8 **trang**, trong đó có **35 câu trắc nghiệm** và **2 câu tự luận**.

A. **TRẮC NGHIỆM (6đ)** Sinh viên chọn **MỘT** câu trả lời đúng nhất cho mỗi câu hỏi trắc nghiệm.

- Mục tiêu của kiến trúc lược đồ 3 lớp (three-schema architecture) là:
  - Giúp người dùng hiểu rõ hơn chi tiết lưu trữ vật lý của cơ sở dữ liệu
  - Tăng cường tính phụ thuộc giữa chương trình và dữ liệu (program-data dependence)
  - Hỗ trợ tính độc lập dữ liệu luận lý và vật lý**
  - Tất cả các câu trên đều sai
- Các bước trong quá trình thiết kế một cơ sở dữ liệu:
  - Thu thập & phân tích yêu cầu, thiết kế ý niệm, thiết kế chức năng, thiết kế luận lý, thiết kế vật lý
  - Thu thập & phân tích yêu cầu, thiết kế ý niệm, thiết kế luận lý, thiết kế vật lý**
  - Thu thập & phân tích yêu cầu, phân tích chức năng, thiết kế ý niệm, thiết kế luận lý, thiết kế vật lý
  - Phân tích chức năng, thiết kế ý niệm, thiết kế luận lý, thiết kế vật lý
- Trong ERD/EERD, phát biểu nào là **ĐÚNG**:
  - Bậc của một kiểu mối liên kết (relationship type) là số lượng các kiểu thực thể (entity type) tham gia vào kiểu mối liên kết đó**
  - Một kiểu thực thể có thể có nhiều thuộc tính khóa nhưng ta chỉ gạch dưới một khóa được chọn làm khóa chính
  - Thuộc tính phức hợp (composite attribute) được biểu diễn bằng hình e-lip và nét đôi
  - Câu A và B đúng
- Trong cách tiếp cận *shared file approach* cho các ứng dụng thì phát biểu nào sau đây là **ĐÚNG**:
  - Vấn đề cấu trúc dữ liệu không uyển chuyển (rigid data structure) xuất hiện và cần phải giải quyết.**
  - Vấn đề dư thừa dữ liệu (Data redundancy) được loại bỏ hoàn toàn.

- Vấn đề truy xuất đồng thời (concurrency control) được hỗ trợ.
- Vấn đề phụ thuộc dữ liệu có tính chất vật lý (physical data dependency) được loại bỏ hoàn toàn.

5. Cho EERD như bên dưới:



Phát biểu nào sau đây là **ĐÚNG**:

- Số sinh viên đại học nhiều hơn số sinh viên cao học..
  - Một sinh viên có thể vừa là sinh viên đại học vừa là sinh viên cao học.
  - Sinh viên đại học có những thuộc tính mà 1 sinh viên có.**
  - Thuộc tính của kiểu thực thể sinh viên bao gồm các thuộc tính của sinh viên đại học và sinh viên cao học.
6. Chọn phát biểu nào **ĐÚNG**:
- Ánh xạ các thuộc tính đa trị luôn tạo ra một quan hệ mới
  - Các thuộc tính dẫn xuất không nên cho phép người dùng nhập dữ liệu trực tiếp
  - Có thể tồn tại một kiểu mối liên kết giữa một kiểu thực thể và một kiểu mối liên kết khác.
  - Câu A và B đều đúng**
  - Cả 3 câu trên đều đúng.
7. Ràng buộc nào có thể bị vi phạm khi thực hiện tác vụ DELETE
- Ràng buộc miền trị (Domain constraint)
  - Ràng buộc tham chiếu (Referential constraint)**

- C. Ràng buộc khóa chính (Key constraint)  
D. Ràng buộc thực thể (entity integrity constraint)  
E. Cả 4 câu trên đều sai
8. Việc thiết kế ERD/EERD nằm ở giai đoạn nào trong quá trình thiết kế một cơ sở dữ liệu:  
A. Thiết kế chức năng (functional design)  
**B. Thiết kế ý niệm (conceptual design)**  
C. Thiết kế luận lý (logical design)  
D. Thiết kế vật lý (physical design)
9. Chọn phát biểu **SAI**: Khái niệm View trong cơ sở dữ liệu quan trọng là vì:  
**A. View có thể cải thiện tốc độ thực thi các câu truy vấn**  
B. View hỗ trợ độc lập dữ liệu (data independence)  
C. View cho phép schema được thay đổi mà không cần phải biên dịch lại các ứng dụng sẵn có.  
D. View hỗ trợ quản lý truy cập (access control) bằng việc cho phép người dùng chỉ được truy suất một phần dữ liệu của cơ sở dữ liệu.
10. Cho quan hệ EMP(ssn, name, dno) và ba câu truy vấn sau:  
Q1: `SELECT name  
FROM EMP  
WHERE dno IN (10, 20);`  
Q2: `SELECT name  
FROM EMP  
WHERE dno = 10 OR dno = 20;`  
Q3: `SELECT name  
FROM EMP  
WHERE dno = 10  
UNION  
SELECT name  
FROM EMP  
WHERE dno = 20;`  
Các câu truy vấn nào tương đương nhau:  
**A. Câu Q1 và Q2**  
B. Câu Q1 và Q3  
C. Câu Q2 và Q3  
D. Cả ba câu Q1, Q2, Q3 tương đương nhau
11. Cơ chế bảo mật trong đó hệ thống tự xác định các hành vi (đọc, ghi) cho phép của một chủ thể (subject) trên một đối tượng dữ liệu dựa sự phân lớp bảo mật của các chủ thể và đối tượng dữ liệu.  
A. Discretionary Access Control (DAC)  
B. Role-Based Access Control (RBAC)  
**C. Mandatory Access Control (MAC)**  
D. Attribute-Based Access Control (ABAC)
12. Nhiệm vụ của người quản trị cơ sở dữ liệu (DBA) là:  
A. Thiết kế cơ sở dữ liệu  
B. Định nghĩa các ràng buộc trong cơ sở dữ liệu

**C. Quản lý việc truy xuất vào cơ sở dữ liệu của người dùng**

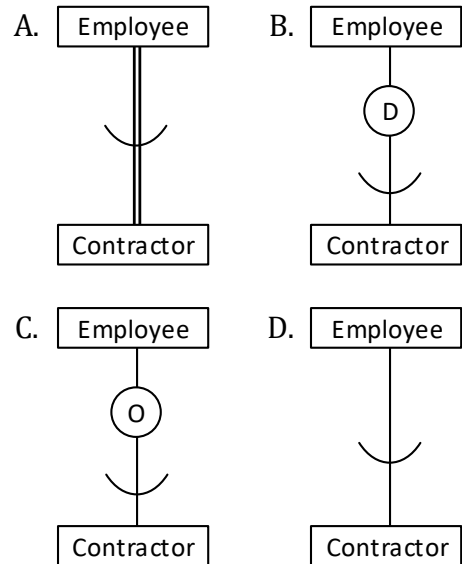
D. Cả 3 câu trên đều đúng

13. Chọn phát biểu **ĐÚNG** về mô hình dữ liệu quan hệ (Relational Data Model)

A. Bậc (degree) của một quan hệ (relation) là số lượng bộ dữ liệu (tuple) có trong quan hệ  
B. Siêu khóa của quan hệ cũng là khóa của quan hệ  
C. Thứ tự các bộ dữ liệu (tuple) là quan trọng và phải khác biệt  
D. Cả 3 câu A, B, C đều đúng

**E. Cả 3 câu A, B, C đều sai**

14. Một công ty cần quản lý nhân viên (*Employee*) của mình. Trong số nhân viên, có những người đặc biệt, làm việc theo dự án, được gọi là *contractor*. ERD/ERRD nào sau đây mô tả tình huống trên:



15. Cho biết một trạng thái (state) của quan hệ R(A, B, C) như sau

A	B	C
a1	b1	c1
a1	b2	c2
a2	b2	c2
a2	b3	c3

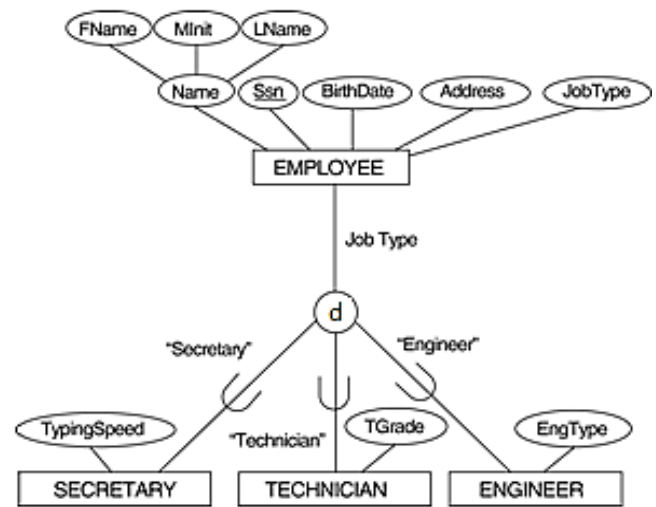
Phát biểu nào sau đây **ĐÚNG**

A. {A, B} là khóa của R  
B. {A, C} là khóa của R  
C. {C} phụ thuộc hàm vào {B}  
**D. {A} không phụ thuộc hàm vào {B, C}**  
E. Tất cả câu trên đều đúng

16. Các bước chính trong quá trình xử lý câu truy vấn viết bởi ngôn ngữ SQL là

- A. Tối ưu hóa câu truy vấn, thực thi câu truy vấn, truy xuất dữ liệu
- B. Phân tích câu truy vấn, tối ưu hóa câu truy vấn, thực thi câu truy vấn
- C. Phân tích cú pháp và chuyển đổi, tối ưu hóa câu truy vấn, sinh code và phiên dịch, thực thi câu truy vấn**
- D. Tất cả đều sai
17. Cơ sở dữ liệu của các hệ thống thông tin phải đối mặt với mỗi nguy hại (threats) nào sau đây:
- A. Mất tính toàn vẹn (loss of integrity)
- B. Mất tính sẵn sàng (loss of availability)
- C. Mất tính bí mật (loss of confidentiality)
- D. Tất cả các câu trên đều đúng**
18. Ưu điểm của mô hình quan hệ thực thể (ER model) là:
- A. Yêu cầu của người dùng được mô tả một cách hình thức và rõ ràng
- B. Tương đối dễ hiểu đối với người dùng thông thường
- C. Mô hình quan hệ thực thể độc lập với các hệ quản trị cơ sở dữ liệu
- D. Tất cả các câu trên đều đúng**
19. Đặc điểm của một quan hệ có cấu trúc tốt (well-structured) là:
- A. Không có dư thừa dữ liệu
- B. Không chứa giá trị NULL
- C. Có một vài bất thường khi người dùng thực hiện thao tác cập nhật dữ liệu
- D. Tất cả các câu trên đều sai**
20. Câu lệnh nào sau đây không gây ra lỗi khi thực thi:
- A. `SELECT DepartmentName, Revenue FROM Department WHERE Revenue = MAX(Revenue);`
- B. `SELECT DepartmentName, Revenue FROM Department WHERE Revenue = (SELECT MAX (Revenue) FROM Department GROUP BY BranchID);`
- C. `SELECT DepartmentName, Revenue FROM Department WHERE Revenue = (SELECT MAX (Revenue) FROM Department);`**
- D. Tất cả các câu trên đều gây ra lỗi

21. Để ánh xạ mối liên kết cha-con (super class – sub class relationship), trong giáo trình có trình bày 4 cách (bước 8). Cho mỗi liên kết cha-con như sau. Xác định các cách ánh xạ phù hợp.



- A. Cách 8A, 8B
- B. Cách 8B, 8D
- C. Cách 8A, 8C, 8D**
- D. Cách 8A, 8B, 8C, 8D
22. Sử dụng cơ sở dữ liệu COMPANY như trong bài học, câu đại số quan hệ nào sau đây trả về nhân viên thuộc phòng "R&D", có làm việc cho một dự án ở Houston và có lương trên \$5000?
- A.  $\sigma_{DNAME='R \& D' \text{ AND } DLOCATION='Houston' \text{ AND } SALARY > 5000}(((EMPLOYEE \bowtie_{DNO=DNUMBER} DEPARTMENT) \bowtie_{SSN=ESSN} WORKS\_ON) \bowtie_{DEPARTMENT.DNUMBER=DEPT\_LOCATIONS.DNUMBER} DEPT\_LOCATIONS)$ .
- B.  $\sigma_{DNAME='R \& D' \text{ AND } PLOCATION='Houston' \text{ AND } SALARY > 5000}(((EMPLOYEE \bowtie_{SSN=MGRSSN} DEPARTMENT) \bowtie_{SSN=ESSN} WORKS\_ON) \bowtie_{PNO=PNUMBER} PROJECT)$ .
- C.  $\sigma_{DNAME='R \& D' \text{ AND } PLOCATION='Houston' \text{ AND } SALARY > 5000}((EMPLOYEE \bowtie_{DNO=DNUMBER} DEPARTMENT) \bowtie_{DNUMBER=DNUM} PROJECT)$ .
- D.  $\sigma_{DNAME='R \& D' \text{ AND } PLOCATION='Houston' \text{ AND } SALARY > 5000}(((EMPLOYEE \bowtie_{DNO=DNUMBER} DEPARTMENT) \bowtie_{SSN=ESSN} WORKS\_ON) \bowtie_{PNO=PNUMBER} PROJECT)$ .**
23. Giả sử tất cả các thuộc tính, các bảng đều hợp lệ, phát biểu nào sau đây gây ra lỗi khi thực thi?
- A. `SELECT * FROM Employee;`
- B. `SELECT LastName, FirstName WHERE EmployeeID = '001';`
- C. `SELECT DeptID, SUM(salary) FROM Employee`

```
WHERE SUM(salary) > 20000
GROUP BY DeptID;
D. SELECT DeptID, SUM(salary)
FROM Employee
GROUP BY DeptID
HAVING SUM(salary) > 100000;
```

E. Cả 2 câu B, C đều gây ra lỗi

24. Sử dụng cơ sở dữ liệu Company như trong bài học, phát biểu nào luôn đúng?

```
Q1: Select E.Fname
      From Employee E S
      Where E.SuperSSN = S.SSN;
Q2: Select E.Fname
      From (Employee E LEFT OUTER JOIN
            Employee S ON E.SuperSSN =
            S.SSN) ;
```

- A. Kết quả của 2 câu Q1 và Q2 giống nhau
- B. Nếu tất cả các hàng trong bảng Employee thỏa E.SuperSSN = NULL, thì kết quả 2 câu Q1 và Q2 giống nhau

C. Nếu có ít nhất 1 hàng trong bảng Employee thỏa E.SuperSSN = NULL thì kết quả của Q1 luôn khác Q2

D. Cả 2 câu A, B đều đúng

25. Kết quả của một phép toán đại số quan hệ giữa 2 bảng Employee (SSN, FName, DNO) và Department (DNO, Dname) được cho như bảng sau. Phép toán đại số quan hệ này có thể là:

SSN	FName	DNO	Dname
Null	Null	100	Research
666	John	200	IT
555	Franklin	300	Finance

A. Phép Outer join

- B. Phép Project
- C. Phép Cross product
- D. Phép Natural join
- E. Tất cả các câu trên đều sai

26. Toàn vẹn dữ liệu (data integrity) là:

- A. Dữ liệu lưu trữ trong cơ sở dữ liệu không bị dư thừa
- B. Dữ liệu lưu trữ trong cơ sở dữ liệu chính xác và nhất quán
- C. Dữ liệu lưu trữ trong cơ sở dữ liệu được bảo mật
- D. Dữ liệu lưu trữ trong cơ sở dữ liệu được chia sẻ đầy đủ
- E. Tất cả câu trên đều sai

27. Cho 2 quan hệ R và S. Biết R có m tuples và n attributes, S có p tuples và q attributes. Phép đại số quan hệ  $R \bowtie P$  có tối đa và tối thiểu bao nhiêu tuple:

- A.  $m + p, \min(n, q)$
- B.  $m + p, |m - p|$
- C.  $mp, 0$
- D.  $mp, m + p$
- E. Tất cả các câu trên đều sai

28. Phép toán đại số quan hệ nào không đòi hỏi các bảng tham gia phải tương hợp kiểu?

- A. Phép hội
- B. Phép giao
- C. Phép hiệu
- D. Phép kết (join)
- E. Các câu trên đều sai

Mô tả sau đây dùng cho các câu trắc nghiệm từ 29 đến 35

Cho kích thước một khối (block) B = 1024 bytes, kích thước con trỏ khối (block pointer) PB = 6 bytes, kích thước con trỏ bản ghi PR = 7 bytes. Một tập tin gồm 50.000 bản ghi (record) có chiều dài cố định, sử dụng cách lưu trữ không phủ (unspanned).

Mỗi bản ghi gồm các trường dữ liệu (field) sau: orderNo (13 bytes), quantity (1 byte), amount (4 bytes), status (1 byte), orderDate (13 bytes), updatedAt (13 bytes), note (35 bytes).

Giả sử tập tin được đánh chỉ mục sơ cấp (primary index) có thứ tự trên trường dữ liệu orderNo.

29. Kích thước của một bản ghi là

- A. 80 bytes
- B. 81 bytes
- C. 86 bytes
- D. 87 bytes
- E. Tất cả đều sai

30. Hệ số phân khối (blocking factor) và số khối (number of file blocks) của tập tin này lần lượt là

- A. 12 và 4167
- B. 13 và 4166
- C. 11 và 625
- D. 625 và 4167
- E. Tất cả đều sai

31. Hệ số phân khối chỉ mục (index blocking factor) ở mức đầu tiên (first level) là

- A. 51
- B. 53

- C. 171  
D. 78  
E. Tất cả đều sai
32. Số lần truy xuất khối trung bình để truy xuất và tìm kiếm một bản ghi với giá trị orderNo cho trước từ tập tin  
A. 78  
B. 79  
C. 7  
**D. 8**  
E. Tất cả đều sai
33. Số lần truy xuất khối trung bình để truy xuất và tìm kiếm một bản ghi có giá trị quantity lớn hơn quantity cho trước từ tập tin  
A. 13  
B. 14  
**C. 2084**  
D. 4167  
E. Tất cả đều sai
34. Giả sử tập tin trên được đánh chỉ mục đa cấp (2 cấp) có thứ tự trên trường dữ liệu orderNo, như vậy số dòng chỉ mục ở cấp thứ 2 (second level) là  
**A. 2**  
B. 3  
C. 78  
D. 79  
E. Tất cả đều sai
35. Giả sử ứng dụng thường xuyên truy xuất các đơn hàng theo ngày tạo đơn hàng (orderDate), loại chỉ mục nào sau đây được tạo thêm để tăng tốc độ tìm kiếm cho các câu truy vấn như trên  
A. Đánh chỉ mục sơ cấp (primary index) trên trường dữ liệu orderDate  
B. Đánh chỉ mục thứ cấp (secondary index) dày (dense) trên trường dữ liệu orderDate  
**C. Đánh chỉ mục thứ cấp (secondary index) thưa (non dense) trên trường dữ liệu orderDate**  
D. Đánh chỉ mục cụm (clustering index) trên trường dữ liệu orderDate  
E. Không thể đánh thêm loại chỉ mục nào nữa vì đã đánh chỉ mục trên trường dữ liệu orderNo

## B. PHẦN TỰ LUẬN (4đ)

**Câu 1 (2.5 đ):** Cho các bảng dữ liệu sau đây

Order(orderNo, orderDate, totalAmount, totalQuantity)

OrderDetail(orderNo, productNo, unitPrice, quantity)

Trong đó orderNo trong bảng OrderDetail là khóa ngoại tham chiếu đến khóa chính orderNo của bảng Order.

Sinh viên thực hiện các yêu cầu sau:

- Viết hàm có input là năm x, output là tổng doanh số bán được trong năm x. (1đ)
- Viết trigger tự động cập nhật totalAmount, totalQuantity của bảng Order khi cập nhật lại số lượng sản phẩm mua được (quantity) trong bảng OrderDetail của cùng một mã đơn hàng (orderNo). (1đ)
- Viết thủ tục có input: tháng x, năm y; thủ tục này xuất ra danh sách gồm các thông tin mã sản phẩm bán được trong tháng x, năm y; tổng số lượng bán được, tổng doanh số bán được của mỗi sản phẩm. (0.5đ)

Sinh viên thực hiện các yêu cầu sau:

- Viết hàm có input là năm x, output là tổng doanh số bán được trong năm x. (1đ)
  - Sinh viên viết đúng cú pháp tạo hàm với input và output đúng (0.25đ)
  - Sinh viên viết đúng câu truy vấn lấy tổng doanh số bán được trong năm x (0.5đ)

**SELECT SUM(totalAmount) FROM Order WHERE EXTRACT (YEAR FROM orderDate) = x**

Lưu ý: hàm SQL Server là YEAR(orderDate) hoặc DATEPART(year, orderDate), MONTH(orderDate) hoặc DATEPART(month, orderDate)

- Sinh viên trả về giá trị trong câu Select vào một biến và ghi được giá trị trả về cho hàm (0.25đ)

- b) Viết trigger tự động cập nhật totalAmount, totalQuantity của bảng Order khi cập nhật lại số lượng sản phẩm mua được (quantity) trong bảng OrderDetail của cùng một mã đơn hàng (orderNo). (1đ)
- Sinh viên tạo đúng trigger theo cú pháp cho hành động **UPDATE** trên thuộc tính **quantity** của bảng **OrderDetail**, chỉ thị **AFTER** (0.25đ)
  - Sinh viên tính đúng totalAmount và totalQuantity cần cập nhật trong bảng Order với đúng mã đơn hàng của hàng dữ liệu orderDetail đang cập nhật (0.75)

*UPDATE Order SET totalQuantity = totalQuantity + (:new.quantity - :old.quantity), totalAmount = totalAmount + (:new.quantity - :old.quantity)\*:old.unitPrice*

*WHERE orderNo = :old.orderNo*

*Lưu ý: với SQL Server thì các giá trị của hàng dữ liệu cũ đang cập nhật lấy từ bảng deleted (ko có bảng updated), và các giá trị mới cập nhật trong bảng dữ liệu inserted*

- c) Viết thủ tục có input: tháng x, năm y; thủ tục này xuất ra danh sách gồm các thông tin mã sản phẩm bán được trong tháng x, năm y; tổng số lượng bán được, tổng doanh số bán được của mỗi sản phẩm. (0.5đ)
- Sinh viên viết đúng câu lệnh tạo thủ tục có 2 input là tháng x kiểu int và năm y kiểu int (0.25)
  - Sinh viên viết đúng câu query lấy danh sách ... (0.25)

*SELECT productNo, SUM(quantity), SUM(unitPrice\*quantity) FROM OrderDetail JOIN Order On OrderDetail.orderNo = Order.orderNo Where Extract (YEAR FROM orderDate) = y and Extract (MONTH FROM orderDate) = x GROUP BY productNo*

**Câu 2 (1.5đ):** Cho mỗi quan hệ CONTRACT như sau

CONTRACT(ContractID, Date, CustomerID, CustomerName, ProductID, ProductName, Quantity, Price)

Và tập phụ thuộc hàm F:

F = {ContractID → (Date, CustomerID, CustomerName),  
CustomerID → CustomerName,  
ProductID → ProductName,  
(ContractID, ProductID) → (Quantity, Price)}

- a) Tìm khóa (key) của CONTRACT. (0.5đ)
- b) Chuẩn hóa mỗi quan hệ trên để đạt được dạng chuẩn 2NF, 3NF (1đ)

*Key = (ContractID, ProductID)*

Chuẩn 2NF

*R1 (ContractID, Date, CustomerID, CustomerName)*

*R2 (ProductID, ProductName)*

*R3 (ProductID, ContractID, Quantity, Price)*

Chuẩn 3NF

*R1 (CustomerID, CustomerName)*

*R2 (ContractID, Date, CustomerID)*

*R3 (ProductID, ProductName)*

*R4 (ProductID, ContractID, Quantity, Price)*