DATABASE QUALIFIER QUESTION, Fall 2007: This question focuses on decomposition and normal forms. There are two parts to the question.

1. The normalization process depends on *factoring* or *decomposing* relations into two or more smaller relationships.

   (1.a) *3 points:*
   One condition on this factoring is that the precise content of the original must be recapturable by joining all of the decomposed parts. *Question: What is this condition called?*

   (1.b) *3 points*
   At what point in the stages of normalization (1NF -> 2NF, 2NF->3NF, 3NF -> BCNF, BCNF -> 4NF, 4NF -> 5 NF) can this condition first occur?

   (1.c) *7 points*
   Provide an example of a decomposition that violates the condition described in 1.a. Explain why it violates the condition. Your example may be for any stage of normalization.

   (1.d) *6 points*
   If a particular relation can be decomposed and then recaptured by joining its parts, can all future contents of the original table also be so decomposed and rejoined? Provide a suitable example.

   (1.e) *6 points*
   If a relation consisting of a three element concatenated key and nothing else has a decomposition that cannot be recaptured by joining its parts, identify at least one way of preventing a lossy decomposition.

2. This question is concerned with integrity conditions that might be applied to decomposed relations.

Given two arbitrary relations:
R1: (A,B, C)
R2: (C,D,E)
that can be naturally joined on C; C is called a *foreign key*.

(2.a) *4 points*
What can happen if there are not matching values for C in both R1 and R2?

(2.b) *6 points*
Show an example of what is described in 2.a.

(2.c) *4 points*
What are the tuples called when there are not matching elements for C in R1 and R2?

(2.d) *5 points*
Various constraints can be applied to avoid missing foreign key values in R1 and/or R2. *Question: What is the "entity integrity" constraint. Provide an example.*

(2.e) *6 points*
Various constraints can be applied to avoid missing foreign key values in R1 and/or R2. *Question: What is the "existence" constraint. Provide an example.*

Database Question 1 **Fall 2007 DQE**

Use the sample relational database table shown in Figure 1 to answer the questions below. The order of tuples in the table is the same as the physical storage order of records. *State any assumptions you make as you solve the problem.*

ssn	fname	lname	license	phone
458-09-1133	Jenny	White	PQ5113	513-867-5309
556-23-9922	Ned	Jones	PQ5114	513-785-9087
789-01-8887	Zelda	Smith	PQ5115	513-556-4471
435-76-5621	Tommy	Greene	PQ5116	513-556-4471
888-90-6430	Ella	Aberdeen	PQ5117	513-785-9087

Figure 1. Sample Relational Database

a. (5 points)
Place a check in the matrix for each of the indexes that could possibly be built over the given field in the table shown in Figure 1. (In other words, if it is impossible to build the given index, do not put a check in the matrix for that entry.)

	primary	clustered	secondary (key)	secondary (non-key)	B+tree
lname					
license					
phone					
ssn					

b. (10 points)
Justify your answer in each of the following cases: (in other words, why did/didn't you place a check in the corresponding box?) Be sure to explain the relevant details of each index type to illustrate your decision.

- license, primary

- ssn, primary

- phone, secondary (non-key)

- phone, secondary (key)

c. (10 points)
Explain how to evaluate the following query using one of the choices you made in the matrix. Identify the choice by field name and index type, e.g., [phone, secondary (key)]. Explain all the steps, including searching/filtering strategy, that are used to evaluate the query.

```
select fname, lname, phone
where fname = “Jenny” and license = “PQ5113”
```

d. (25 points)
For the physical layout of the data file, assume a blocking factor of two records per block as shown in Figure 2. Using this data file, draw a valid index structure for (i) a primary index, (ii) a secondary (non-key), and (iii) a B+ tree from the matrix for one of the fields. Identify the choice by field name and index type as in part (c).

block 1:	458-09-1133	Jenny	White	PQ5113	513-867-5309
	556-23-9922	Ned	Jones	PQ5114	513-785-9087
block 2:	789-01-8887	Zelda	Smith	PQ5115	513-556-4471
	435-76-5621	Tommy	Greene	PQ5116	513-556-4471
block 3:	888-90-6430	Ella	Aberdeen	PQ5117	513-785-9087

Figure 2. Data File Blocks