Beckham Carver

20 February 2023

COSC 4820

Dr. Buckner

Homework 5

**Question 1: [15 points]**

Let R(A, B, C, D, E) be decomposed into relations with following sets of attributes: {A, B, C}, {B, C, D}, and {A, C, E}. For each of the following sets of FD’s, use the chase test to tell whether the decomposition of R is lossless. For those that are not lossless, give an example of an instance of R that returns more than R when projected on the decomposed relations and rejoined. (Please note that an instance of a relation is one in which there is actual data.)

1. [5 points] AC → E and BC → D.

Using the FD’s AC->E and BC->D we can unsubscript e1 and d1 respectively, we end our chase test with the final tableau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** |
| a | b | c | d | e |
| a2 | b | c | d | e2 |
| a | b3 | c | d3 | e |

Because our tableau contains a row with no subscripts this decomposition of R is lossless.

1. [5 points] A → D, D → E, and B → D.

Using the FD’s B->D, A->D, and D->E we can unsubscript d1 d3 and e1 respectively, we end our chase test with the final tableau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** |
| a | b | c | d | e |
| a2 | b | c | d | e2 |
| a | b3 | c | d | e |

Because our tableau contains a row with no subscripts this decomposition of R is lossless.

1. [5 points] A → D, CD → E, and E → D.

Using the given FD’s A → D, CD → E, and E → D we can unscubscript/modify d1 and e1 respectively, we end out chase test with the final tableu:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** |
| a | b | c | d1 | e |
| a2 | b | c | d | e2 |
| a | b3 | c | d1 | e |

Because this final tableau does not contain a fully unsubscripted row, this decomposition is lossy.

As an example, below is a instance of R that returns additional data to R when decomposed and rejoined:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **OS** | **Version** | **Bug** | **Month** |
| M1 | Catalina | 3.0 | Seg fault | Jan |
| M2 | Catalina | 3.0 | Null Data | Mar |
| M1 | Big Sur | 2.2 | Seg fault | Jan |

Decomposition into {ABC}, {BCD}, and {ACE}

|  |  |  |
| --- | --- | --- |
| **Model** | **OS** | **Version** |
| M1 | Catalina | 3.0 |
| M2 | Catalina | 3.0 |
| M1 | Big Sur | 2.2 |

|  |  |  |
| --- | --- | --- |
| **OS** | **Version** | **Bug** |
| Catalina | 3.0 | Seg fault |
| Catalina | 3.0 | Null Data |
| Big Sur | 2.2 | Seg fault |

|  |  |  |
| --- | --- | --- |
| **Model** | **Version** | **Month** |
| M1 | 3.0 | Jan |
| M2 | 3.0 | Mar |
| M1 | 2.2 | Jan |

Rejoining the decompositions we get the final table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **OS** | **Version** | **Bug** | **Month** |
| M1 | Catalina | 3.0 | Seg fault | Jan |
| M1 | Catalina | 3.0 | Null Data | Jan |
| M2 | Catalina | 3.0 | Null Data | Mar |
| M2 | Catalina | 3.0 | Seg fault | Mar |
| M1 | Big Sur | 2.2 | Seg fault | Jan |

Our final table is contains additional information than what we began with in R, therefore the decomposition is lossy.

**Question 2: [15 points]**

For each of the sets of FD’s in Question 1, are the dependencies preserved by the decomposition? A yes or no answer is sufficient. Please refer to the discussion in Section 3.4.4. (5 points each)

1. Yes, dependencies are preserved
2. No, dependencies are not preserved
3. No, dependencies are not preserved

**Question 3: [40 points]**

For the following we will look at 3NF violations and subsequent decomposition of the  
relations.

**(a)** Indicate all of the 3NF violations, if none, so state.

***i)*** [2 points] R(A, B, C, D) with FD’s B → C and B → D.

B->C and B->D are both 3NF violations because their left-hand sides are not keys/superkeys, nor are they trivial FD’s, and their righthand sides do not compose a key.

***ii)*** [2 points] R(A, B, C, D) with FD’s AB → C, BC → D, CD → A, and AD →B.

No 3NF violations.

***iii)*** [2 points] R(A, B, C, D, E) with FD’s AB → C and DE → C, and B → D.

AB->C, DE->C and B->D are all 3NF violations because their left-hand sides are not keys/superkeys, nor are they trivial FD’s, and their right hand sides do not compose a key.

***iv)*** [2 points] R(A, B, C, D, E) with FD’s AB → C and C → D, D → B, and D → E.

D->E is the only violation because {AB} is a key, {D} and {B} are members of a key, but for D->E, D is not a key nor is E a member of a key.

**(b)** Decompose the above relations, as necessary, into collections of relations that are  
in 3NF and project the FD’s onto the new relations.

***i)*** {B->C, B->D} is our minimal basis, first decompose into R1(B,C) with {B->C} and R2(B,D) with {B->D}, none are super keys therefore, add relation R3(A,B). Now all three relations R1, R2, and R3 are in 3NF.

***ii)*** Relation is already in 3NF.

***iii)*** {AB->C, C->D, D->B, D->E} is our minimal basis, first decompose into R1(A,B,C) with {AB->C} and R2(D,E,C) with {DE->C} and R3(D,B) with {D->B}. Because none are a superkey we add R4(A,B,E). Now all four relations are in 3NF.

***iv)*** {AB->C, C->D, D->B, D->E} is our minimal basis, first decompose into R1(A,B,C) with {AB->C} and R2(C,D) with {C->D} and R3(D,B) with {D->B} and R4(D,E) with {D->E}. Because R1 is a superkey for our relation we are done.

**Question 4: [20 points]**

Consider the relation Courses(C, T, H, R, S, G), whose attributes may be thought of informally as course, teacher, hour, room, student, and grade. Let the set of FD’s for Courses be C → T, HR → C, HT → R, HS → R, and CS → G. Intuitively, the first says that a course has a unique teacher, and the second says that only one course can meet in a given room at a given hour. The third says that a teacher can be in only one room at a given hour, and the fourth says the same about students. The last says that students get only one grade in a course.

1. [5 points] What are all the keys for Courses?

The minimal key for Courses is {HS}, only larger superkeys can be constructed beyond HS.

1. [5 points] Verify that the given FD’s are their own minimal basis. (You must demonstrate this not just make a flat statement.)

The definition of a minimal basis is to: have a singleton right side, removing an FD would make the basis no longer equivalent to the original set, and removing an attribute from an FD would make it no longer equivalent to the original set.

All of the given FD’s have a singleton right side. However both HS and HT can determine R and HT is not a key nor does it’s right side compose a key; therefore the given FD’s are not a minimal basis because we can remove HT->R.

1. [5 points] Use the 3NF synthesis algorithm to find a lossless-join, dependency preserving decomposition of R into 3NF relations.

The 3NF violations are  C->T, CS->G, HR->C and HT->R. Using the minimal basis {C->T, HR->C, HS->R, and CS->G} we can decompose into R1(C, T) with {C->T} and R2(C, G, S) with {CS->G} and R3(H, R, S) with {HS->R} and R4(H, R, T) having {HT->R}. Because we have a key {HS} we are done. These relations are in 3NF.

1. [5 points] Are any of the relations from (c) not in BCNF? Explain why or why not.

All relations in C are in BCNF firstly they are in 3NF from the synthesis algorithm. To be in BCNF one of the conditions must be met: the right side is a subset of the left, OR the left side is a key/superkey.

For every relation in C the FD’s are a key for their respective relations.s