Lab7-design-normalzation 1NF

**Purpose:**

* To begin the **normalization** of user views from **UNF** to **1NF**
* How to create the **Unnormalized Form** (UNF) from a user view
* How to identify **Repeating Groups** (multivalued dependencies)
* How to create the **First Normal Form** (1NF) from the UNF.

**Definition: Normalization is the process of assigning attributes to entities in such a way that data redundancies are reduced or eliminated.**

**Definition: User Views can be individual descriptions, reports, forms, or lists of data that are required to support the operations of a particular database user.**

**Part A. Creating the UNF from a User View. The goal here is to create a single entity from the data found in the user view.**

**Method:**

–**Examine the user view**

–**Identify attributes**

–**Describe the entity using DBDL**

1.

Here is an example of a user view. This report lists the customers for the Premiere Corporation.

**Premiere Corporation  
Customer List**

**Cust Name Street City State Zip Sales**

**No Code Rep**

# Num

124 Sally Adams 482 Oak Lansing MI 49224 03

256 Ann Samuels 215 Pete Grant MI 49219 06

311 Don Charles 48 College Ira MI 49034 12

315 Tom Daniels 914 Cherry Kent MI 48391 06

405 Al Williams 519 Watson Grant MI 49219 12

412 Sally Adams 16 Elm Lansing MI 49224 03

522 Mary Nelson 108 Pine Ada MI 49441 12

567 Tran Dinh 808 Ridge Harper MI 48241 06

587 Mara Galvez 512 Pine Ada MI 49441 06

622 Dan Martin 419 Chip Grant MI 49219 03

\*\*\*\*\*\*\*\* End of report \*\*\*\*\*\*\*\*

**Step 1: examine the report:**

As you examine this report, you can see that it contains a title, a line of column headings and the body of the report. Each line in the body of the report contains data about a particular customer. For example we can see that in the first line of the body of the report, there is data about Customer 124. This customer’s name is Sally Adams and she lives at 482 Oak in Lansing, Michigan (MI). The Sales Representative who calls on Sally Adams is Sales Rep Number 03.

**Step 2: Identify attributes:**

The attributes (or characteristics) of a User View can often be found by simply looking at the column headings. In our Customer Report we see that we have the following attributes: Customer Number, Customer Name, Street, City, State, Zip Code and Sales Representative Number.

**Step 3: Describe the entity using DBDL:**

Database Design Language (DBDL) is a standardized way of describing entities of a relational database.

1. Choose an appropriate name for the entity. We chose to name our entity **Customer** because each line in the report describes customer data.
2. List the attributes you found in the user view inside square brackets, giving each attribute a suitable attribute name.
3. Determine which attribute would be suitable as a primary key and underline that attribute.

The DBDL for the entity resulting from our Customer user view would look as follows:

**Customer [ CustNo, CustName, CustStreet, CustCity, CustSt, CustZip, CustRep ]**

**NOTE: it should be noted that *calculated fields or derived fields* should not be included in the DBDL!! There is no calculated field in the above example, but there will be one in a later example.**

2. Now you try it. Examine the following report:

**Premiere Corporation**

**Parts List**

**Part Part Qty Class Warehouse Price**

**Number Description On Hand of Part**

AX12 Iron 104 HW 3 24.95

AZ52 Dartboard 20 SG 2 12.95

BA74 Basketball 40 SG 1 29.95

BH22 Cornpopper 95 HW 3 24.95

BT04 GasGrill 11 AP 2 149.99

BZ66 Washer 52 AP 3 399.99

CA14 Griddle 78 HW 3 39.99

CB03 Bike 44 SG 1 299.99

CX11 Blender 112 HW 3 22.95

CZ81 Treadmill 68 SG 2 349.95

\*\*\*\*\*\*\*\* End of report \*\*\*\*\*\*\*\*

What type of data does each line in the report represent?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

What attributes can you identify from the user view?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

What would be a suitable name for the UNF entity? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Which attribute would be suitable as a primary key? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Describe the UNF entity using DBDL:

**NOTE: Enter this UNF DBDL on the lab submission page which is the last page of this lab**.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Part B. Recognizing Repeating Groups**

**Definition: Unnormalized form (UNF) is a table that contains one or more *repeating groups.***

**Definition: A repeating group is an attribute or collection of attributes within a table that has multiple values for a single occurrence of the primary key for that table.**

For example, looking at the following User View, we see that for each Customer number, it is possible to have **multiple values** for the order number and order date attributes. Looking at customer 124, you can see that there are 2 orders for that customer (12489 and 12500). The same is also true for customer 522. In a larger system customers will place a lot of orders. Again, for 1 occurrence of customer there can be many occurrences of order and order date, meaning the order number and the date of the order represent a repeating group.

**Premiere Corporation**

**Customer Orders**

**Customer Name Order Order**

**Number Number Date**

124 Sally Adams 12489 2001-09-02

12500 2001-09-05

256 Ann Samuels 12495 2001-09-04

311 Don Charles 12491 2001-09-02

315 Tom Daniels 12494 2001-09-04

522 Mary Nelson 12498 2001-09-05

12504 2001-09-05

\*\*\*\*\*\*\*\* End of report \*\*\*\*\*\*\*\*

We identify repeating groups in our DBDL by placing brackets around them. For example the DBDL for this User View would look something like this:

**Customer [ CustNo, CustName, ( OrderNum, OrderDate ) ]**

Notice the brackets around the OrderNum and Orderdate attributes. This quickly and easily identifies a repeating group to someone who is reading the DBDL.

**Common Mistake:** A common mistake that students make is to incorrectly identify repeating data as a repeating group. For example, the previous report could also have been printed in the following way:

**Premiere Corporation**

**Customer Orders**

**Customer Name Order Order**

**Number Number Date**

124 Sally Adams 12489 2001-09-02

**124 Sally Adams**  12500 2001-09-05

256 Ann Samuels 12495 2001-09-04

311 Don Charles 12491 2001-09-02

315 Tom Daniels 12494 2001-09-04

522 Mary Nelson 12498 2001-09-05

**522 Mary Nelson**  12504 2001-09-05

\*\*\*\*\*\*\*\* End of report \*\*\*\*\*\*\*\*

The fact that the Customer Number and Name for Sally Adams and Mary Nelson have been repeated on multiple lines does **not** make it a repeating group!! You should still identify that for one customer number, there are multiple values for the order number and date. Note that this does not mean that EVERY customer number will have multiple order numbers, just that this user view shows that it is possible for some customer numbers to have multiple values for these attributes. Therefore the repeating group must be identified in the DBDL. Note also that it is possible to have more than 1 repeating groups.

**Part C. Creating the 1NF from the UNF**

**Definition: A table is in 1NF if it does not contain any repeating groups.**

**Therefore, the process of taking an entity from UNF to 1NF, involves removing the repeating groups.**

**Method:**

* **Rewrite the DBDL for the UNF table without the repeating group.**
* **Choose a primary key for the repeating group.**
* **Create an additional table for the repeating group. The primary key for this additional table will be the concatenation of the primary key of the original table PLUS the primary key you have chosen for the repeating group.**

For example, using our previous report from part B, we had the following:

UNF: Customer [ CustNo, CustName, ( OrderNum, OrderDate ) ]

**Step 1: Rewrite the UNF without the repeating group:**

Customer [ CustNo, CustName ]

**Step 2: Choose a primary key for the repeating group.**

**OrderNum** would make a suitable primary key for the repeating group as it uniquely identifies the data in the repeating group.

**Step 3: Create an additional table for the repeating group with a concatenated key (original key + key of repeating group)**

CustOrder [ CustNo, OrderNum, OrderDate ]

**Note that I have called this additional table CustOrder to show that the table is a combination of customer and order data.**

So the **1NF** of our user view is as follows:

Customer [ CustNo, CustName ]

CustOrder [ CustNo, OrderNum, OrderDate ]

**NOTE: If the UNF does not contain any repeating groups, it is already in 1NF. In other words, the UNF and 1NF will be exactly the same!**

Additional Note: Although what you have done above is correct some of you should be able to see that in the Customer order table or the table that contains orders, only the Order Number needs to be the primary key. For 1 order number there can only be 1 customer number and 1 date. At this stage it will not matter which result you have. The above result will rectify itself in 2NF

**Write the answer of following questions and show it personally next week to get marks during lab period.**

**1. Answer from Part A:**

**UNF:**

**2. For the following User View, determine the UNF and the 1NF and hand in this page to your instructor.**

**Premiere Corporation**

**Order Detail Report**

**Order Order Cust Part Part Number Quoted Total**

**Number Date Number Number Desc Ordered Price**

12489 2001-09-02 124 AX12 Iron 11 14.95 164.45

12491 2001-09-02 311 BT04 GasGrill 1 149.99 149.99

BZ66 Washer 1 399.99 399.99

12494 2001-09-04 315 CB03 Bike 4 279.99 1,119.96

12495 2001-09-04 256 CX11 Blender 2 22.95 45.90

12498 2001-09-05 522 AZ52 Dartboard 2 12.95 25.90

BA74 Basketball 4 24.95 99.80

12500 2001-09-05 124 BT04 GasGrill 1 149.99 149.99

12504 2001-09-05 522 CZ81 Treadmill 2 325.99 651.98

\*\*\*\*\*\*\*\* End of report \*\*\*\*\*\*\*\*

**UNF:**

**1NF:**