ECU178 Computer Science: 220CT Data and Information Retrieval Coursework

Due on March 13th 2015

Robert Rigler : 4939377

Contents

Task 1 Database Design

This task in Database Design consists of four activities. The first three involve normalising the given data to third normal form, and the fourth is to produce and Entity Relationship diagram of the normalised relations. For each activity I will give a detailed step by step explanation of how I completed each activity.

Activity 1: First Normal Form

To put this data into First Normal Form (1NF), I need to:

1. Identify any repeating redundant data and remove it from the current Entity
2. Place the data into a new Entity
3. Create a relationship with a primary key from one Entity as a foreign key in the other.

Step 1: Identify Redundancy

On inspecting the data, I can see that there are multiple instances of repeating data.
Orders' ID: *CON-2237, CON-2356* and *CON-1234* all have repeating data entries for fields: *Equipment, Qty,* and *Unit Price* .

Step 2: Create New Entity

Removing the *Equipment, Qty,* and *Unit Price* fields and placing them in a new entity, leaves me with two entities as shown below.

Step 3: Relationships and Keys

To complete the First Normal Form, a relationship needs to be created between the entities.
I created the relationship by including the *Order ID* attribute as a foreign key in the *ItemOrder* entity. *Order ID* is used as a Primary Key for the *Order* entity. In the *ItemOrder* entity, no one attribute can be used to uniquely identify a single record. For this reason, I have created a concatenated key using the attributes *Equipment* and *Order ID*. The concatenated key can now be used to uniquely identify each record.

1NF : Diagram

Order	(Order ID
	Supplier ID
	Client Name
	Client Address
	Date
	Total Price)

ItemOrder	(*Order ID
	Equipment
	Qty)
	Unit Price)

1NF : Data

Order entity

Order ID	Supplier ID	Client Name	Client Address	Date	Total Price
CON-2237	168	Coventry Building Services Ltd	Units 2-4, Binley Industrial Estate, CV3 2WL	14-Dec-14	£99.00
CON-3664	527	Allied Construction Ltd	34,Lythalls La Industrial Estate, CV6 6RG	16-Jan-15	£36.00
CON-2356	169	Rioh Builds Ltd	Unit 12,Stoneleigh Park, CV8 2UV	12-Feb-15	£280.00
CON-1234	032	Grand Design Ltd	32-34, Bilton Industrial Estate, CV3 5YB	16-Apr-15	£23.00

ItemOrder entity

Order ID	Equipment	Qty	Unit Price
CON-2237	Butterfly Valve	2	£5.00
CON-2237	3/4" Locknut	6	£1.50
CON-2237	Sch 40 Blk Pipe	4	£20.00
CON-3664	Thin Stranded Copper Wire	6	£6.00
CON-2356	Sch 40 Blk Pipe	3	£20.00
CON-2356	4x8x3/4 Cos Plywood	2	£10.00
CON-2356	3/4" EMT	2	£50.00
CON-2356	Duplex Ivy Rec	1	£100.00
CON-1234	Sch 40 Blk Pipe	1	£20.00
CON-1234	3/4" Locknut	2	£1.50

Activity 2: Second Normal Form

To put this data into the Second Normal Form(2NF) I need to ensure that the attributes are completely dependant on the primary key, i.e., that no attribute is only dependant on one part of the primary key. This can be done in two steps:

1. Test each attribute for complete dependency on the primary key.
2. Remove any partially dependent attributes to a new entity and assign a primary key.

For this particular set of data, the *Order* entity does not have a concatenated key and therefore is already in the second normal form.

Step 1: Testing each attribute.

Primary Key	Attribute	Functionally Dependant?
Order ID, Equipment	Qty	Yes, dependant on both
Order ID, Equipment	Unit Price	No, dependant on Equipment only

Step 2: Entity, Relationship and Key

From my testing, I found that the attribute *Unit Price* is not functionally dependant as it is only dependent on *Equipment*, but not *Order ID*.
I moved *Unit Price* into a new entity called *Item* and then I created an relationship between the *Item* and *ItemOrder* entities by repeating the *Equipment* attribute in the *Item* entity. *Equipment* bceomes the primary key.

Entities in Second Normal Form: Diagram

Order	(Order ID	**Item**	(Equipment	**ItemOrder**	(*Order ID
	Suppler ID		Unit Price)		*Equipment
	Client Name				Qty)
	Client Address				
	Date				
	Total Price)				

Entities in Second Normal Form: With data

Order entity

Order ID	Supplier ID	Client Name	Client Address	Date	Total Price
CON-2237	168	Coventry Building Services Ltd	Units 2-4, Binley Industrial Estate, CV3 2WL	14-Dec-14	£99.00
CON-3664	527	Allied Construction Ltd	34,Lythalls La Industrial Estate, CV6 6RG	16-Jan-15	£36.00
CON-2356	169	Rioh Builds Ltd	Unit 12,Stoneleigh Park, CV8 2UV	12-Feb-15	£280.00
CON-1234	032	Grand Design Ltd	32-34, Bilton Industrial Estate, CV3 5YB	16-Apr-15	£23.00

Item entity

Equipment	Unit Price
3/4" EMT	50.00
3/4" Locknut	1.50
4x8x3/4 Cos Plywood	10.00
Butterfly Valve	5.00
Duplex Ivy Rec	100.00
Sch 40 Blk Pipe	20.00
Thin Stranded Copper Wire	6.00

ItemOrder entity

Order ID	Equipment	Qty	Unit Price
CON-2237	Butterfly Valve	2	£5.00
CON-2237	3/4" Locknut	6	£1.50
CON-2237	Sch 40 Blk Pipe	4	£20.00
CON-3664	Thin Stranded Copper Wire	6	£6.00
CON-2356	Sch 40 Blk Pipe	3	£20.00
CON-2356	4x8x3/4 Cos Plywood	2	£10.00
CON-2356	3/4" EMT	2	£50.00
CON-2356	Duplex Ivy Rec	1	£100.00
CON-1234	Sch 40 Blk Pipe	1	£20.00
CON-1234	3/4" Locknut	2	£1.50

Task 1 Database Design continued on next page...

Activity 3: Third Normal Form

To place the data into Third Normal Form 3NF I need to ensure that all attributes are only dependent on the Primary Key, and not Non-Key Attributes. This can be achieved in two steps:

1. Test each attribute for dependency on the primary key.
2. Remove all transitive dependencies to a new entity with the correct primary key and relationship.

For this particular set of data, both the *Item* and *ItemOrder* entities are already in the Third Normal Form.

Step 1: Testing for Transitive Dependency

Order Entity

Primary Key	**Attribute**	**Transitive Dependency?**
Order ID	Supplier ID	Yes: Supplier ID can be found if we know Client Name or Address
Order ID	Client Name	Yes: Client Name can be found if we know Supplier ID or Address
Order ID	Client Address	Yes: Client Address can be found if we know Client Name or Supplier ID
Order ID	Date	No: Only dependent on Primary Key
Order ID	Total Cost	No: Only dependent on Primary Key

Using this table I have identified that *Supplier ID, Client Name,* and *Client Address* all have transitive dependencies, and need to be moved to a new entity.

Step 2: Entity, Relationship and Key

I created a new entity called *Customer*, and moved the three attributes with transitive dependencies into it. I then made *Supplier ID* the Primary Key and created a relationship between the *Customer* and *Order* entities by repeating the *Supplier ID* attribute as a foreign key in the *Order* entity.

Entities in Third Normal Form: Diagram

Order (Order ID
*Suppler ID
Total Price) Date

Item (Equipment
Unit Price)

ItemOrder (*Order ID
*Equipment
Qty)

Customer (Supplier ID
Client Name
Client Address)