University of Toronto
CSC343, Fall 2016

Assignment 3

Due: Wednesday, Wed, December 7, at 8:00 pm sharp!

IMPORTANT: Read the Piazza discussion board for any updates regarding this assignment. We will provide a summary of key clarifications, and it is required reading. Check it regularly for updates.

Learning Goals

By the end of this assignment you should have:

1. become fluent in ER Model;

2. learned how to reason about data design problems and tradeoffs; and

3. learned how to formally reason about functional dependencies and how they are used in data design.

4. gained experience in designing a database schema, from conceptual design through to DDL.

Part 1: Conceptual Design

For this part of the assignment, you will design an Entity-Relationship (ER) schema for the manufacturing company application described below. Your design should include all information needed to support the processes and documents of the company that are described. If a process, such as payroll, is not described, then you should not model information needed to support that process.

Specify all entity types and all attributes of each entity type. All attributes will be assumed to be single-valued (meaning the maximum cardinality constraint is 1) unless you place a different cardinality constraint on the attribute.

Indicate key attributes and all cardinality and existence constraints on the relationship types (using the min-max notation used in class). All of this information should be clearly indicated on the diagram. You should use the diagram conventions from class.

Hand in the following files:

1. closets.pdf: A pdf file containing an ER diagram depicting your schema. It must show a complete, readable diagram, on one page. Handwritten schemas will not be accepted. You may use any diagraming package of your choice (powerpoint is fine).

2. closets.txt: A plain text file containing a translation of your ER diagram into a relational schema. Please express the relational schema in SQL DDL (complete with constraints) that runs in Postgres on the CS Teaching Labs. If there are any constraints in the ER diagram that cannot be expressed in the Postgres SQL DDL, state what they are and why they cannot be expressed. You can do this in comments, using the "--" syntax.

3. demo.txt: A plain text file containing an interaction with the Postgres shell, on the CS Teaching Labs, that shows you importing your DDL file.

Application Description

Note: Below are many details about the company, and some terminology about manufacturing that may be new to you. As we are learning in class, part of your job is to learn the terminology (be prepared to look up some

definitions) and to identify which of these details are relevant to the application described. You may need to seek some clarification from the client (your profs are playing that role), but do expect to first read this a few times and invest some time digesting it.

ClosetClosetClosets is a small manufacturer of closets. The company has a number of product lines ranging from their basic storage units and bookcases, to wall-closets, and luxury walk-in closets. ClosetClosetClosets markets its products to a wide variety of retail outlets. At the present time, the company has about 50 employees, and sales have been increasing by about 25% per year.

Production Process   Production is mostly "to stock", rather than to order. Production starts with the purchase of seasoned rough-cut lumber of various species and grades. Most of the lumber is purchased locally and there is negligible lead time for most items. The lumber is first sent to the rough mill, where the basic parts used in the closet are cut and prepared. The rough mill consists of power saws, planers, gluing machines, and other tools that are used to cut and form these parts. A typical production order requires about one week in the rough mill (this includes queue time, when the order is waiting for other orders to be completed). From the rough mill, an order is sent to the finish mill, which is responsible for boring, routing, and sanding the rough-milled parts. A typical order requires one week in this department. The order is then moved to the Assembly Department, which assembles the finished parts into end items of a closet. Finally, these items travel to the Finish Department, which applies coats of stain and lacquer. The items are then inspected and packed for shipment. A typical order requires one week each in assembly and finishing. Thus the total time required to complete a typical production order is four weeks.

Organization Chart   An organization chart is shown below. The president, Knotts, started the company and is the principal owner. Three functional managers report to Ms. Knotts: Sheth (finance and administration), Kim (manufacturing) and Wang (marketing). In turn, manufacturing is divided into four areas: fabrication (which includes the rough and finish mills), assembly, finishing and materials. The materials manager is responsible for production planning and inventory control. The company stresses quality in its products, and all employees are expected to inspect their work and maintain high quality standards. In the following sections, the key documents and processes of the company are described.
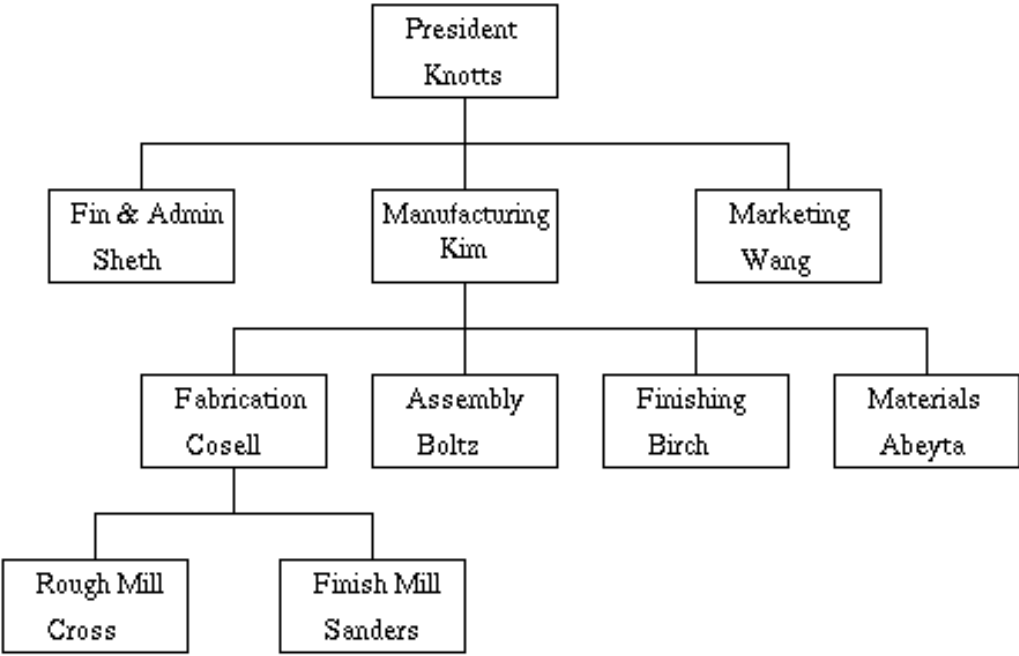

Figure 1: Current Organization Chart

Order Processing   There are about 500 customers at the present time. Most of these customers are retail closet stores, including some large chains. Orders are transmitted by mail or taken by telephone. At the present time, the average order volume is about 50 orders per day. On the average, two orders are outstanding at a given time for each customer. An order may specify from one to twelve items; at the present time, each order specifies an average of six items.

Orders are taken by a clerk in the marketing department. When an order is taken, a new (unused) order number is created. This along with the date, customer number, name and billing address are placed on the order form. The customer specifies the product number and quantity desired for each product in the order. The marketing clerk uses the product number to retrieve the product name, description and unit price of the product associated with this number from the system. After verifying with the customer that this is the correct product, all this information is entered on the order form.

When an order is taken for a new customer, a new customer record is created and stored. The clerk uses a forms interface to obtain a new, unique customer number and enters the name and billing address of the customer. The system stores information about all customers including their year-to-date purchase total, their credit limit, outstanding balance and the negotiated customer discount percentage. The marketing manager negotiates the credit limit and discount with a customer.

The system verifies new orders before they are filled. If the amount of an order, combined with a customer's outstanding balance exceeds the customer's credit limit, then a notice to this effect is sent to the customer and the new order is not processed until the customer makes sufficient payments.

Materials   The materials manager maintains the information about each product, including the quantity on hand, and the "reorder point". When the available stock falls below the reorder point, a new production order is released to obtain a specified reorder quantity (or lot size) of this item. In addition, for each product, the total quantity of production orders in progress is maintained. At present, there are about 200 products representing ten product lines. The materials manager determines the product number for each new product. The product number includes two parts: a code for the product line and a number that is unique within a specific product line.

Billing   After an order has been approved, invoices are prepared. An invoice has an invoice number, the customer information (name and address), the order number and date of the invoice. Each product ordered is listed with its number, description, unit price, and order quantity. The total bill per item and for the whole invoice is shown on the invoice. If a customer has negotiated a discount, then the discounted amount is applied to the invoice balance and the total amount due is calculated as the total price minus the discount amount and minus any payments.

For each order, invoices are consecutively numbered starting with one. If full payment is received within a month, no further invoices are sent. Otherwise, a new invoice, with the new amount due is created. All payments made before the invoice date are taken into account when calculating this new amount.

Bill of Materials   A bill of materials (BOM) is a listing of the parts (or components) required to build a product or subassembly. The figure below contains an example bill of materials for a simple bookcase. A bookcase consists of two sides, three shelves, a back, and a panel. It is assembled using 2-inch wood screws. While bookcases are relatively simple, more complex products are possible that require 20 or more parts. In addition, each part could have its own BOM. So both products and parts may be composed of a set of parts. On average, a product requires about ten different parts. The parts are either stocked or produced as needed. A given part may be used in several different products.

Bill of Materials

Product/Part Number: _M128_   Product Description: _Bookcase_

PartNo	Description	Quantity	Location	Code
139409	Side	2	A	1
328179	Shelf	3	A	1
421835	Back	1	B	1
218367	Panel	1	A	1
98210	Screw, 2in	24	C	2

Figure 2: Example Bill of Material

Routing   A Routing is a list of the operations (in the sequence they are performed) required to produce a product. Thus, there is one Routing for each product. The Routing for the bookcase is shown in the figure below. Only products (not parts) have routings. Operation numbers may be reused (only) across different products.

Routing

Product Number: _M128_   Product Description: _Bookcase_

OpNo	Operation Description	Workcenter	Setup Time	Std Oper Time
010	Cut Sides (20	Radials	0.20	0.10
020	Cut Shelves (3)	Radials	0.20	0.10
030	Cut Back	Radials	0.25	0.15
040	Cut Panel	Radials	0.10	0.05
050	Groove Sides (2)	Routers	0.20	0.05
060	Sand Parts	Sanders	0.00	0.20
070	Assemble	Assembly	0.20	0.15
080	Finish	Finishing	0.20	0.50

Figure 3: Current Organization Chart

Work Centers   Each operation in a routing is performed at a work center. There are currently ten work centers, each with a specific capacity.

Production   When the inventory level of a product falls below the reorder point, the materials manager issues a production order to replenish the stock. A production order for a product has a job number, a quantity and an estimated completion date. In addition the production order is assigned to a work center for coordination. This work center may not be assigned to perform all operations on the product's routing, but the work center manager is in charge of ensuring the production order is filled on time. Each production order pertains to only one product. However, there may be more than one order in progress for a given product at a particular time. Typically, about ten production orders are in progress at a given time.

Part 2: : Functional Dependencies, Decompositions, Normal Forms

Show all of your steps so that we can give part marks where appropriate. There are no marks for simply a correct answer. And you must follow all instructions concerning alphabetical ordering for full marks.

1. Consider a relation schema R with attributes $ABCDEFGHIJ$ with functional dependencies S.
   $$S = \{A \rightarrow E, BC \rightarrow AE, C \rightarrow ACF, DE \rightarrow A, EFG \rightarrow AB, I \rightarrow J, J \rightarrow AI\}$$

   (a) State which of the given FDs violate BCNF.

   (b) Decompose R into BCNF using a lossless-join decomposition into a set of relations in BCNF. You must follow the BCNF decomposition algorithm given in class. Make sure it is clear which relations are in the final decomposition, and don't forget to project the dependencies onto each relation in that final decomposition. Because there are choice points in the algorithm, there may be more than one correct answer. List the final relations in alphabetical order (order the attributes alphabetically within a relation, and order the relations alphabetically).

   (c) State whether your decomposition is dependency preserving. If not, state which FDs are lost.

2. Consider a relation P with attributes $ABCDEFGH$ and functional dependencies T.
   $$T = \{A \rightarrow B, BC \rightarrow ACE, C \rightarrow B, EF \rightarrow CG, EFG \rightarrow ABCD, GH \rightarrow ABCD\}$$

   (a) Compute all minimal keys for P.

   (b) Compute a minimal basis for T. In your final answer, put the FDs into alphabetical order. Within a single FD, this means stating an FD as $XY \rightarrow A$, not as $YX \rightarrow A$. Also, list the FDs in alphabetical order ascending according to the left-hand side, then by the right-hand side. This means, $WX \rightarrow A$ comes before $WXZ \rightarrow A$ which comes before $WXZ \rightarrow B$.

   (c) Using your minimal basis from the last subquestion, employ the 3NF synthesis algorithm to obtain a lossless and dependency-preserving decomposition of relation R into a collection of relations that are in 3NF. Do not "over normalize", this means that you should combine all FDs with the same left-hand side into a single relation.

   (d) Does your schema allow redundancy? Explain why or why not.

Show all of your steps so that we can give part marks where appropriate. There are no marks for simply a correct answer.

What to hand in for Part 2

Type your answers up using LaTeX or Word. Hand in your typed answers, in a single pdf file called part2.pdf.

Marking

Part 1 60 points. Part 2 40 points.

Some parting advice

It will be tempting to divide the assignment up with your partner. Remember that both of you probably want to answer all questions the final exam.