Lab 2: Database requirements analysis and design

Due date: Tuesday, April 16, midnight.

Note: This lab involves a number of deliverables that are due before April 16. The due date above is for the final set of the deliverables.

This is a **team lab**. Each team submits one set of deliverables. Each team member receives the same grade for the assignment.

Course Project Information

This quarter, CPE 366 comes with a group course project. General information about the project is provided below.

Customer. Our customer for this quarter is David Cumberland, VP, Engineering of Shopatron.com, an e-commerce company based in San Luis Obispo, and a Cal Poly CSC department graduate.

Project. The customer asks you to design a database system supporting inventory management operations. Shopatron has recently added inventory management as an option offered to its customers in conjunction with its primary services — processing and mediating orders for products placed from manufacturer's web sites (and fulfilled by retailers). Inventory management is a separate system within Shopatron's infrastructure. Each team will duplicate a large portion of the actual functionality of this system.

Lab Overview

In the course of this lab, you will

- receive an overview of the customer requirements via direct communication with the customer (April 4);

- prepare an initial design of the database for the project (April 9);
- elicit additional requirements for the database via direct communication with the customer (April 11);
- extend the database design to model the customer requirements in full (April 16).

The lab assignment spans four lab periods, and has four distinct steps, each of which serves a different goal and yields a set of deliverables. This document specifies your assignment for each of the periods as well as final deliverable.

The overview of the lab structure is below.

April 4, 2013: 9:40 - 11:00am: Groups are announced. Dave Cumberland joins us for an initial presentation. He will discuss his company, its business model and will give a brief and broad overview of the desired database and software application.

April 9, 2013: 9:40 - 11:00am: Initial database design. We will work on identifying major components of the database model and on discovering the key relationships between them.

April 11, 2013: 9:40 - 11:00am: Q&A session with the customer. Each team will have an allocated portion of time for questions. Questions/answers must be recorded and posted later.

April 11 - April 16, 2013: teams work on the full database design.

Assignment

April 4: Group Creation and Initial Presentation

We will spend the first few minutes of the lab period on April 4 on group creation. Group lineup will be announced by the instructor at the beginning of the lab. We will ensure that the group lineups are feasible (e.g., that we are not missing three people on a single 4-person team). Teams will contain four to five people.

Following team formation, our customer, Dave Cumberland, will present an overview of your project, followed by a short, informal Q&A session. The main purpose of the presentation is to give you a good understanding of what the project entails. After the presentation, we will release customer-supplied documentation:

- Database description/requirements.
- Software use cases.

These two documents, combined with the data, which will be released upon your completion of Lab 2 will serve as the basis for your project.

Assignment. Lab attendance is required for everyone - we will be organizing teams based on who comes to the lab. Please don't forget your nametags (I will bring the name tags for those who do not have them yet). During the presentation, you are welcome to take notes - those will come in handy later. Upon receiving customer documentation, study it in preparation for the April 9 lab session.

Deliverables. Each team shall come up with a name. Further, you shall create a team wiki page on the `csc366 wiki`. The page, at the very least shall contain team name and names and emails (Cal Poly userIds are fine) for all team members.

April 9: Initial Design

We will build the database model in two stages. First, *on April 9*, you will work out the overall database model, and will determine the most important components of the model. After that, on April 11-16 you will refine the model to incorporate specific requirements that the customer placed on the identified components.

Assignment: Each team shall fully participate in the model-building activities that occur during the *April 9* lab period. As the result of these activities, each team should produce a draft initial (simplified) Entity-Relationship model for the customer database.

Deliverables: The key deliverable of this stage is the simplified E-R model of the customer database. Each team will produce a draft of the model by the end of the *April 9* lab period. Each team then has until **Wednesday, April 10, 11:59pm** to submit the finalized model. The model shall be put on the csc366 wiki. The model description shall consist of the following:

1. List of entity sets. For each entity set specify its attributes, and the primary key. Identify weak (if any) entity sets, specify their owners.
2. List of relationship sets. For each relationship set, identify participating entity sets, list any attributes, identify its type (one-to-one, one-to-many, many-to-many), and whether it is the defining relationship set for a weak entity set.
3. An E-R diagram of the model.

The first two parts of the E-R model description should be a text document (plain text, Word, Postscript, PDF are all ok formats). The E-R diagram shall be drawn using some graphics editor or presentation program. I use `xfig` to create diagrams that are embedded into latex/PostScript files. I use MS PowerPoint to create diagrams taht are embedded into MS Word documents.

Feel free to embed the diagram with the rest of the model description, or to put as a standalone.

Bring and submit the **hardcopy** of your model description to the **Tuesday, April 14, 2009** lab.

April 11: Question and Answer Session

The initial design that you complete by *April 10* will make simplifying assumptions about the database. Portions of you initial E-R model will have to go through thorough revision and extension before you obtain the real E-R model for the customer database. The process of refining your model will be done in two steps. The first step is a Q&A session with the customer, during which each team gets a chance to ask questions about the project. The second step is the actual mdoel refinement process that will take place after the Q&A session.

Preparation. Before the *April 11* lab, each team shall study the materials provided by the customer. The goal of each team is to identify any aspects of the application domain, information about objects to be stored in the database, customer needs (w.r.t. the database), *that require clarification.* Each team shall prepare a list of questions that the team wants to ask the customer during the Q&A session. Note, that in addition to the customer, you may direct your questions at the instructor.

Please note: initial list of questions **must be prepared** by each team independently of other teams. However, once lists are prepared, I allow for teams to compare them. I expect that many questions will coincide or be similar (it is ok if multiple groups have similar/same questions). It is also ok for one group to come up with a question, no other group has asked — there is no need for other groups to add it to their lists, since each question only needs to be asked once, and all teams benefit from the answer.

Each team shall post its list of questions to the csc366 wiki (inside the team's workspace).

Q&A session. Each team will be given 7-10 minutes of time to ask their questions and receive answers from the customer and the instructor. We will set the order of questioning by a simple lottery at the beginning of the lab. I strongly suggest that each team sits and works together during the Q&A. Please note the following:

- Listen carefully to the questions other groups are asking. It is very likely, that some of your team's questions will be posed by teams who get to ask questions earlier.
- If your group runs out of questions yield your time to the next group.
- If your group has unanswered questions left after its Q&A period, wait until all groups ask their questions. If the question has still not been asked, you will have a chance to ask it at the end of the lab.
- You are certainly allowed to ask questions that are not on your list - some questions might occur to you as a result of things said during the Q&A session.

Each team is responsible for recording the answers to all questions its members have asked (whether those questions were prepared or were asked on the spot).

After each team had a chance to ask questions, if any additional questions remain, a free-form Q&A period will begin and will last until the end of the lab period. Any questions that went unasked, or unanswered during the Q&A session should be submitted in writing (via email) to the instructor. The questions will be forwarded to the customer, who, in turn, will provide answers prior to the Tuesday, April 16 lab.

Deliverables. After the Q&A session, all teams, in collaboration shall create the project knowledge base as part of the csc366 wiki. The structure of the knowledge base is left up to the teams (generally speaking, you want to break the entire knowledge base by topic: e.g., group information about one type of data in one place, and information about another type of data — in another). At the "bottom" of the knowledge base hierarchy shall lie wiki pages that contain questions that the teams asked and answers received from the customer (plus any additional information that teams think may be useful). A single question can appear in multiple places in the knowledge base (as it may be applicable to more than one "leaf" topic).

The knowledge base is one of the two **main deliverables** for this lab. The initial state of the knowledge base shall be completed by the main Lab 2 due date. Since the knowledge base is part of the csc366 wiki, feel free to maintain and update it throughout the project. The intent is to have a useful and accessible resource for your further development.

April 11- April 16: Database Design

Assignment. The main task of this lab is for each team to prepare an initial full database design document. By the end of the *April 14* lab session, each team shall have the completed simple design, and shall have an understanding of which parts of the simple design need to be revised, enhanced and extended. Between April 14 and April 16, each team shall work on the appropriate revisions, enhancements and extensions.

During April 16 lab session, each will will work with the instructor on the full database model. Depending on the level of progress we may conduct one or two structured activities at the beginning of the lab period, or, the entire lab period may be given to the teams for team meeting activities, with instructor participating in each team's meeting in turn.

It is expected that after the April 16 lab, each team will have a conceptual model of the database for the project.

Deliverable. Each team is given time until **Tuesday, April 16, midnight** to finalize and submit the database model document.

The design document prepared and submitted by each team shall, at a minimum, contain the following information:

- List of entity sets for the proposed database.
- List of attributes for each proposed entity set.
- Identification of primary keys for each proposed entity set.
- Identification of all weak entity sets in the proposed database, and of their discriminating attributes.
- Identification of any other entity set constraints for the proposed entity sets.
- List of relationship sets for the proposed database. For each relationship set, the following must be indicated:
  - All participating enitity sets.
  - All clarifying relationship set attributes.
  - Type of the relationship set (one-to-one, one-to-many, many-to-many).
  - If the relationship set is an identifying one for a weak entity set.
- List of other relationship set constraints.
- List of any class hierarchies and/or aggregates in the database model.
- List of relationship sets (with all the information as above) associated with the aggregates.

- E-R diagram of the proposed database design.

The design document must be typeset (handwritten submissions will not be accepted). See comments above about the software to use to draw the E-R diagram. The E-R diagram shall contain all entity sets and relationship sets. It shall also contain all key attributes for all entity sets and all relationship set attributes. Other attributes can be omitted from the diagram for clarity. All constraints that can be shown on the diagram, shall be shown.

The design document shall begin with the name of the group, and the list of group members. Note, that in general, **all documents** submitted by each team during the course of the project must contain the team name and the list of students.

Deliverables Overview and Submission Instructions

Overall, the lab has the following formal deliverables:

No.	Deliverable	Due:	Softcopy	Hardcopy
1	Simplified Database Design	April 10, midnight	Yes	Yes No
2	List of questions for Q&A session	April 11, 4:40pm	Yes	No
3	Project knowledge base	April 16, midnight	Yes	No
4.	Database Design document	April 16, midnight	Yes	No

Submission. Softcopy submissions must appear on your team's wiki page by the due date/time. Your team's contribution to the project knowledge base shall be up on the wiki by the due date/time.

Due to printing limitations in the labs, no hardcopy submissions are required.

Please, keep the soft copies of all submitted documents. You will be working with them in the labs that follow.

Grading. Due to my travel schedule (I am out of town April 17, and April 19 - April 21), I will not be able to grade your submissions until after the weekend, so you will receive my comments on April 23 during the lab period. Lab 3 will start on April 18; I expect you to take my comments into account when revising your Lab 3 solutions on April 23.