

Introduction to Database Systems

Programming Assignment 1

Submission Deadline: October 12, 2021 before 23:59

Assignment 1: Database Design

Objectives

The purpose of this activity is to practice the fundamentals of Database Design using ERDs, Relational Data Model and SQL.

Submission Requirements

- Please submit your results/output in a PDF file. Submit your code in a text (.txt) file.
- The name of the files must be your student number such as 100131001.pdf or 100131001.txt.
- File must be uploaded to the A01 submission link provided. The name of submission on the link must also be your student number.
- Your diagrams can be hand drawn or digital. In both cases, the images must be clear.
- If the files do not open properly or the content is not clear, then you will be awarded zero.
- **Deadline is October 12, 2021 before 23:59. Late submission is not permitted under any circumstances.**
- **This is an individual assignment.**
- **Your submissions will be verified using Turnitin (or some other suitable tool) for originality. 60% or more similarity will be awarded zero in the assignment and reported to the department. We may report similarity less than 60% if it is of significant nature.**

Note: In database design, variations are always possible. If the business requirements are taken care of in a proper logical way as per the data modeling guidelines, the answer will be considered correct.

Warning: Please add the necessary headings and labels to your report so that the TAs can understand the different parts properly. Anything that we cannot understand will be awarded zero.

General Instructions


- Please make sure that your document is easy to understand; clearly add the question number, part number, captions and foot notes wherever required. If the TA can't locate the answer, then it is your responsibility.
- Make sure that images/screenshots are clear. If the image is big then split it into multiple parts. Clearly write their purpose. You can add multiple images even if the question statement doesn't say so to make sure that your answer is easy to comprehend.
- Highlight the significant parts of each image so that TA can easily identify the required answer.
- Add necessary explanation to make sure that TA can understand different parts of your document.

Question 1: ER Modeling

[5 Marks]

Solve Q 4.17 from the book and add the answer to your pdf submission file. Below is the statement for your reference. You need to draw/design only the additions/modifications. Clearly indicate what changes you introduced and their purpose. **The figure number mentioned in the problem statement is wrong and the bank schema is given in figure 3.22.:**

Figure 3.22

4.17. Consider the BANK ER schema in ~~Figure 3.21~~ , and suppose that it is necessary to keep track of different types of ACCOUNTS (SAVINGS_ACCTS, CHECKING_ACCTS, ...) and LOANS (CAR_LOANS, HOME_LOANS, ...). Suppose that it is also desirable to keep track of each ACCOUNT's TRANSACTIONS (deposits, withdrawals, checks, ...) and each LOAN's PAYMENTS; both of these include the amount, date, and time. Modify the BANK schema, using ER and EER concepts of specialization and generalization. State any assumptions you make about the additional requirements.

Question 2: Reverse Engineering: Relational Model to ER Model

[10 Marks]

Call for Papers or CFP is a popular term in academic research. Fundamentally, it is an announcement for the researchers to submit their research articles for upcoming conferences, journals or books. The CFPs can be classified into these three categories, for: Conferences, Journals or Books.

We want to design a database to record all the CFPs. Below are the links for three events for your understanding (one example from each kind of CFP):

- Conference CFP: <http://cs-conferences.acadiau.ca/ant-21/#callforPapers>
- Journal CFP: <https://www.journals.elsevier.com/journal-of-information-security-and-applications/call-for-papers/special-issue-multimedia-forensics-and-data-hiding-recent-advances>
- Book CFP: See email in Appendix A

For this system, consider the Relational Model given in the form of SQL code/script in the file **a01script.sql**. Draw an ER Diagram that represents this Relational schema and add the diagram to your submission document.

Question 3: Creating Schema

[4 + 2 + 6 Marks]

Use the SQL script (**a01script.sql**) provided with the assignment to prepare the database schema. Before you begin, you must make sure that you don't have schema objects with the same names already in your database. Furthermore, you have to make the following additions to the schema provided in this script:

- Create an additional summary table "CFPCount" which stores the count of CFPs of each type. Also, write a suitable trigger to update this table "CFPCount" whenever a new CFP row/entry is inserted or an existing is deleted. The details of this implementation are up to your imagination; as long as it fulfills the purpose, your solution will be considered correct.

You are required to:

- Submit all the code for schema creation, after your additions, in a script file. Name this file as yourID_q3.sql.txt.
- Run this script on your MySQL server and add the screenshots. Your screenshots must:

- Show the list of tables (there should be none) before the script is executed; use some suitable technique to do that.
- Show the statement that executes the script and the success status of the script.
- Show the list of tables and other schema objects after the script is executed.
- The whole script must execute successfully.

Question 4: Scraping and Inserting Data

[2 + 6 Marks]

Our next problem in this assignment is to extract the data from the CFP announcements and then insert it into the Database using the suitable INSERT statements. **Each student must process at least ONE CFP announcement and create the INSERT statements for this CFPs.** Select ONE (or more from the additional file in case you are interested to insert more data) CFP announcement(s) **from your designated file (Appendix C)**, scrape the data from the CFP(s) (use the method described in Appendix B) and prepare the INSERT statements accordingly. **While scraping the data, you are not required to be precise; make sure you process some meaningful information.** At least 1 Topic, 1 People and 1 Activity must be inserted for a CFP. You are required to:

- Submit all the code in a script file. Name this file as yourID_q4.sql.txt.
- Run this script on MySQL server and add the screenshots. Your screenshots must:
 - Show the statement to execute script and the success status of script.
 - The whole script must execute successfully.

Appendix A

[TCCC-ANNOUNCE] Handbook on Blockchain: CALL FOR CHAPTERS

1 message

D. A. T. <tran.umass@gmail.com>
Reply-To: "D. A. T." <tran.umass@gmail.com>
To: tccc-announce@comsoc.org

Wed, Sep 2, 2020 at 11:23 PM

Dear Colleagues,

-- Please help us forward to anyone who might be interested --

CALL FOR BOOK CHAPTERS

Book Title: HANDBOOK ON BLOCKCHAIN

To be published by SPRINGER NATURE (Series "Springer Optimization and Its Applications": <https://www.springer.com/series/7393>)

Important Dates:

Submission deadline: November 1, 2020

Authors' notification: January 1, 2021

Revisions due: March 1, 2021

Camera ready version due: April 1, 2021

Tentative publication date: July 2021

Introduction:

Blockchain enables a decentralized digital society where people can contribute, collaborate, and transact without having to second-guess trust and transparency. It is the distributed ledger technology behind the success of Bitcoin, Ethereum, and many emerging applications and platforms that disrupt various sectors, including finance, education, healthcare, environment, transportation, government, to name a few.

We are calling for chapter contributions for the Handbook on Blockchain, to be published by Springer Nature in 2021. The book will be a one-stop source of reference that covers both theoretical and application aspects of blockchain, with the following key features:

- 1) Presents the history and evolution of blockchain from its preliminary concepts to fundamental theories that lay the foundation for blockchain
- 2) Presents advanced techniques, architectures, and concepts aimed to optimize the core operations of a blockchain network, making it more efficient and scalable
- 3) Presents innovative ways to enable next-generation blockchain networks so that they become adoptable for more real-world applications
- 4) Presents how blockchain can offer enhanced benefits when integrated with other technological systems involving Cloud/Edge Computing, Internet of Things, Big Data, and Artificial Intelligence
- 5) Presents a comprehensive coverage of real-world applications of blockchain, touching many industries including finance, education, healthcare, environment, logistics, and smart cities.

Topics of Interest:

We especially encourage chapter contributions that are of survey or expository forms. The detailed list of topics includes but is not limited to:

I. Blockchain Basics

- 1) History of Blockchain
- 2) Bitcoin Theory: Structure and Operation
- 3) Blockchain Architecture and Networking
- 4) Cryptography Algorithms
- 5) Mining and Consensus
- 6) Ethereum Platform and Smart Contracts
- 7) Enterprise-level Blockchain: Hyperledger Network
- 8) Distributed applications (dAPPS)
- 9) Decentralized Autonomous Organizations (DAOs)

II. Blockchain Optimization and Advances

- 1) Blockchain Security
- 2) Blockchain Consensus
- 3) Blockchain Scalability
- 4) Blockchain Privacy
- 5) Blockchain Modeling and Analysis
- 6) Inter-Blockchain Services
- 7) Data Science for Blockchain

- 8) Blockchain Visualization
- 9) Next-generation Blockchain
- 10) Blockchain and Cybersecurity
- 11) Blockchain and Artificial Intelligence
- 12) Blockchain and Data Services
- 13) Blockchain and Internet of Things
- 14) Blockchain and Mobile Computing

III. Blockchain Applications

- 1) Blockchain for Banking and Finance
- 2) Blockchain for Supply Chain
- 3) Blockchain for Education
- 4) Blockchain for Smart Cities
- 5) Blockchain for Consumer Products and Retail
- 6) Blockchain for Government
- 7) Blockchain for Automotive
- 8) Blockchain for Healthcare
- 9) Blockchain for Travel and Transportation
- 10) Blockchain for Energy
- 11) Blockchain for Agriculture
- 12) Blockchain for Insurance
- 13) Blockchain for Real Estate
- 14) Blockchain for Charity

Submission Guidelines:

All book chapters need to be formatted according to Springer's Book Manuscript instruction:

<https://www.springer.com/de/authors-editors/book-authors-editors/resources-guidelines/book-manuscript-guidelines/manuscript-preparation/5636>

Minimum length requirement: 30 pages for each chapter

All chapter authors (corresponding authors, co-authors, contributing authors) should include their affiliation and email address in their manuscript. Their email addresses (also published in the chapters) will be used by Springer to provide authors with a personal MySpringer account where they can download a free copy of the eBook. In addition, all authors will be offered a 40% discount on any eBook or print book order from our web shop. Both the free eBook and the author discount will be activated on MySpringer.

Authors should obtain permissions from the previous publisher for any material they want to reuse.

Please email your book chapter to the editors (contact information below).

Handbook Editors:

 Professor Duc A. Tran, PhD
 Department of Computer Science
 University of Massachusetts, Boston (USA)
duc.tran@umb.edu

Professor My T. Thai, PhD
 Department of Computer and Information Sciences and Engineering
 University of Florida (USA)
mythai@cise.ufl.edu

IEEE Communications Society Tech. Committee on Computer Communications

<http://committees.comsoc.org/tccc/>

TCCC Announce: For announcements concerning computer networking and communications.

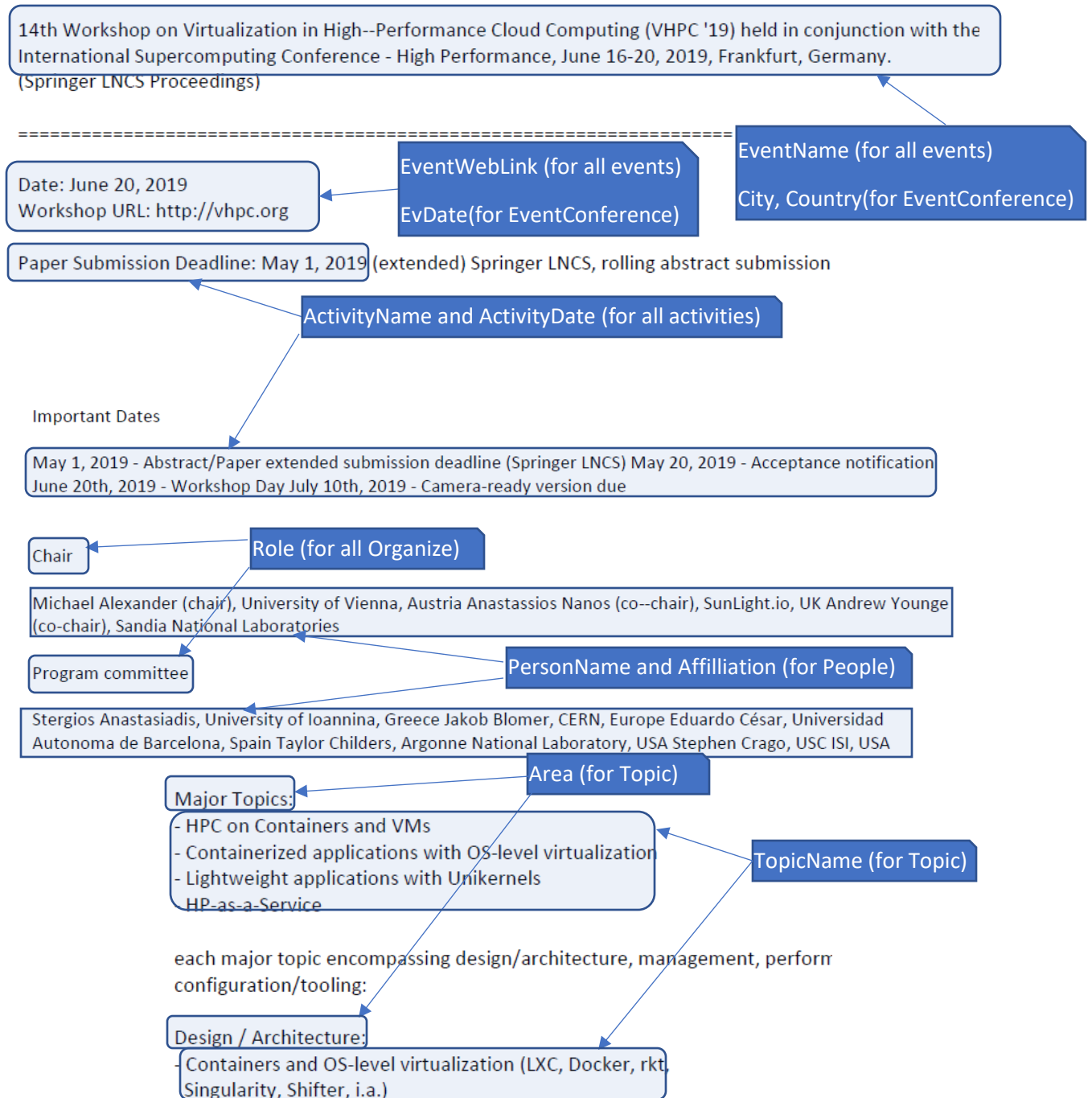
tccc-announce@comsoc.org

<https://comsoc-listserv.ieee.org/>

Appendix B: How to locate data in CFPs

There might be some minor variation in the terminology used in the different CFPs. When looking for People who Organize a certain CFP, you must extract at least one person from the CFP. Similarly, for Topic and Activity, you must extract at least one topic for that CFP and one activity.

The names highlighted refer to the field names used in script provided for Question 1.



For "Topic", if you can't locate the "Area" then make it same as "TopicName".

The CFPs for the “EventConference” usually have phrases such as Conference, Symposium, Workshop in their name/title:

14th Workshop on Virtualization in High--Performance Cloud Computing (VHPC '19) held in conjunction with the International Supercomputing Conference - High Performance, June 16-20, 2019, Frankfurt, Germany.
(Springer LNCS Proceedings)

Fourteenth International Conference on Digital Information Management (ICDIM 2019) August 22-24, 2019 Irish Computer Society, Dublin, Ireland
(www.icdim.org)

International Symposium on High Performance Mobile Computing & Wireless Networks for HPC (MCWN 2019) As part of The 17th International Conference on High Performance Computing & Simulation (HPCS 2019)

15 - 19 July 2019 Dublin Ireland

The CFPs for the “EventJournal” will always have the name of Journal mentioned and usually have the phrase “Special Issue” in the name/title:

Call for papers

IEEE Internet of Things Journal Special Issue on Emerging Trends and Challenges in Fog Computing for IoT

The CFPs for the “EventBook” will always have the phrase “Call for Book Chapters” or “Call for Chapters” in the name/title:

Call for Book Chapters

BOOK TITLE: ***Internet of Things and Secure Smart Environments: Success and Pitfalls***

Chapman & Hall/CRC Big Data Series

To be published by CRC Press, Taylor & Francis, USA

URL: <https://sites.google.com/view/iot-smart-env>

Appendix C: Data files to scrape data

In each row, the left column has the student IDs who are going to use the file from the corresponding right cell for scraping the data. In the order of your ID, process ONE CFP from the corresponding PDF file. For example, student with ID “11158966” will use first CFP from the file “CFPs--1.pdf”; student with ID “31130763” will use last(20th) CFPs from the file “CFPs--5.pdf”.

Each file is for 12 students. Tenth file is extra and has some more CFPs; in case you want to process extra or replace the one assigned to you, pick one from here.

Student ID (combined for both sections, sorted in ascending order)	Group	File to Use
11158966, 11175987, 11229853, 11306168, 11356325, 11358014, 11358383, 11365720, 11374760, 11376691, 11391127, 11396853	1	CFPs--1.pdf
11489630, 11490843, 11492424, 11494492, 11499381, 11500056, 11500307, 11502071, 11510657, 11512856, 11513580, 11514954	2	CFPs--2.pdf
21552897, 21555114, 21558245, 21569467, 21566306, 21538900, 21565823, 21565358, 21568444, 21568694, 21573866, 21535368	3	CFPs--3.pdf
21503876, 21591575, 21592765, 21602188, 21644562, 21900045, 21909301, 21913665, 21917021, 21913779, 21914936, 21910160	4	CFPs--4.pdf
31130763, 31132080, 31134452, 31345829, 31334789, 31336799, 31339019, 31339542, 31339812, 31340292, 31341067, 31345707	5	CFPs--5.pdf
31660005, 31669723, 31669406, 31669613, 31602367, 31604819, 31672589, 31678051, 31689166, 31678399, 31614352, 31627107	6	CFPs-6.pdf
41870297, 41940413, 41940942, 41957433, 41958571, 41986432, 47003187, 47041575, 47047754, 47106956, 47121823, 47129230	7	CFPs-7.pdf
47280041, 47281031, 47303652, 47305764, 57346016, 57353129, 57367137, 57374448, 57395609, 67397829, 67405085, 67488651	8	CFPs-8.pdf
67602111, 67612847, 67637711, 77643453, 77646803, 77766098, 77796665, 77847146, 77935305, 78036897, 78185041, 78256743	9	CFPs-9.pdf
Some Extra CFPs		CFPs--Extra.pdf

Each CFP starts with an email header as following:



From: VHPC 19 <vhpc.dist@GMAIL.COM>
Sent: April 3, 2019 12:00 PM
To: tccc-announce@COMSOC.ORG
Subject: [TCCC-ANNOUNCE] CFP VHPC19: HPC Virtualization-Containers: Paper due May 1, 2019 (extended)

Your designated CFP will be one of the three kinds; you are required to process only ONE CFP, not one from each kind.