



Xi'an Jiaotong-Liverpool University

西交利物浦大學

XJTLU Entrepreneur College (Taicang) Cover Sheet

Module code and Title	Database Development and Design (DTS207TC)	
School Title	School of AI and Advanced Computing	
Assignment Title	002: Assessment Task 2 (CW)	
Submission Deadline	23:59, 29th Dec (Friday)	
Final Word Count	NA	
If you agree to let the university use your work anonymously for teaching and learning purposes, please type "yes" here.	Yes	

I certify that I have read and understood the University's Policy for dealing with Plagiarism, Collusion and the Fabrication of Data (available on Learning Mall Online). With reference to this policy I certify that:

- My work does not contain any instances of plagiarism and/or collusion.
- My work does not contain any fabricated data.

By uploading my assignment onto Learning Mall Online, I formally declare that all of the above information is true to the best of my knowledge and belief.

Scoring – For Tutor Use					
Student ID					
Stage of Marking	Marker Code	Learning Outcomes Achieved (F/P/M/D) (please modify as appropriate)			Final Score
		A	B	C	
1 st Marker – red pen					
Moderation – green pen	IM Initials	The original mark has been accepted by the moderator (please circle as appropriate):			Y / N
		Data entry and score calculation have been checked by another tutor (please circle):			Y
2 nd Marker if needed – green pen					
For Academic Office Use		Possible Academic Infringement (please tick as appropriate)			
Date Received	Days late	Late Penalty	<input type="checkbox"/> Category A <input type="checkbox"/> Category B <input type="checkbox"/> Category C <input type="checkbox"/> Category D <input type="checkbox"/> Category E		Total Academic Infringement Penalty (A,B, C, D, E, Please modify where necessary) _____

Students



Xi'an Jiaotong-Liverpool University

西交利物浦大學

The assignment must be typed in an MS Word document and submitted via Learning Mall Online to the correct drop box. Only electronic submission is accepted and no hard copy submission.

All students must download their file and check that it is viewable after submission. Documents may become corrupted during the uploading process (e.g. due to slow internet connections). However, students themselves are responsible for submitting a functional and correct file for assessments.

Assignment Project Quiz Exam Essay Help
WeChat: cestbon_688
Email: accoder_overseas@163.com



Database Development and Design (DTS207TC)

Assessment 002: Individual Coursework

Due: Dec 29th, 2023 @ 23:59

Weight: 40%

Maximum Marks: 100

Overview & Outcomes

This assignment aims to gain experience in understanding the internal functionality of different database management systems, including RDBMS, XML, Object-relational databases, and gain experience in designing data warehouses to perform OLAP operations. The coursework will be assessed for the following learning outcomes:

- A. Identify and apply the principles underpinning transaction management within DBMS.
- B. Demonstrate an understanding of advanced SQL topics.
- C. Illustrate the issues related to Web technologies and DBMS and XML as a semi-structured data representation formalism.
- D. Identify the principles underlying object-relational models.
- E. State the main concepts in data warehousing and data mining.

Assessment Tasks

Your task is to answer every question by carefully reading the questions and guidelines for the system setup. Record your thoughts and assumptions, where necessary, while reporting your answers. There are the following five parts for this coursework corresponding to five learning outcomes of this module.

1. DBMS Indexing
2. Transaction Management
3. XML Data Modeling
4. Object-Relational Database
5. Data Warehousing and OLAP

Marking Criteria

The coursework will be graded out of 100 marks with a 40% weightage of the final grade. Your final report should be a complete, polished artefact that incorporates all the necessary detail from each of the components. This is an opportunity for you to pull all of your work from the term together into one complete project.

There are five (5) parts and twenty (20) marks available for each part. Marks will be awarded based on the level of correctness of each answer. For example, 100% for fully correct with required explanation/justification, partial marks based on the level of incorrectness or missing required details.



1: DBMS Indexing (20 Marks)

Consider a disk with block size $B = 512$ bytes. A block pointer is $P = 6$ bytes long, and a record pointer is $PR = 7$ bytes long. A file has $r = 40,000$ CUSTOMER records of fixed length. Each record has the following fields: Name (30 bytes), ID (10 bytes), email (10 bytes), Address (40 bytes), Phone (10 bytes), and Birth_date (10 bytes). (4*5=20 marks.)

Q1(a) What would be the number of blocks required to store this data file assuming an unspanned file organization?

Q1(b) How many blocks are required to search for a record if the data file is ordered on ID as a key field? Compare with the average line search performance of unordered file organization.

Q1(c) Suppose the file is ordered by the key field ID, and we have a primary index on ID. How many blocks are required to search and retrieve a record from the data file given its ID value using the primary index?

Q1(d) Suppose we want to construct a multi-level primary index. Calculate:

- How many index entries and the number of blocks are required for each level of the multi-level index? For example, the number of first-level index entries and the number of first-level index blocks, number of second-level index entries and number of second-level index blocks and so on.
- What would be the maximum number of blocks required to search for and retrieve a record from the data file with the multi-level index on the ID field?

Q1(e) Suppose we create a B+ index on ID, and the tree order is 50. What is the maximum capacity of the B+ tree of height 4, and how many blocks accesses are needed to search for and retrieve a record from the file—given its ID value—using the 4 level B+ tree index?

2: Transaction Management (20 Marks)

Consider a database with objects x , y , and z and three concurrent transactions $T1$, $T2$, and $T3$ (given below), which perform read(r) and write(w) operations on database objects. We use the notation from the lecture, where $r_i(x,y,z)$ and $w_i(x,y,z)$ mean that transaction T_i reads and writes objects x,y,z , respectively.

Transactions	
T1	$r1(x); r1(z); w1(x)$
T2	$r2(z); r2(y); w2(z); w2(y)$
T3	$r3(x); r3(y); w3(y)$

Below, two schedules, $S1$ and $S2$, are given. Your task is to answer the questions given for each schedule:



S1		
T1	T2	T3
r1(x)	r2(z)	r3(x) r3(y)
r1(z)		
w1(x)		
	r2(y) w2(z) w2(y)	w3(y)

S2		
T1	T2	T3
r1(x)	r2(z)	r3(x)
r1(z)	r2(y)	r3(y)
w1(x)	w2(z)	w3(y)
	w2(y)	

Q2(a) Consider the schedule S1 to answer the following questions:

- Draw the serializability (precedence) graphs for schedule S1 based on conflicting statements in T1, T2, and T3. **(4 marks)**
- State whether the schedule is conflict serializable (conflict-free) or not. State which statement of which transaction is conflicting with other transactions. **(4 marks)**
- If the schedule is serializable, write down the equivalent serial schedule. **(2 mark)**

Q2(b) Consider the schedule S2 to answer the below questions:

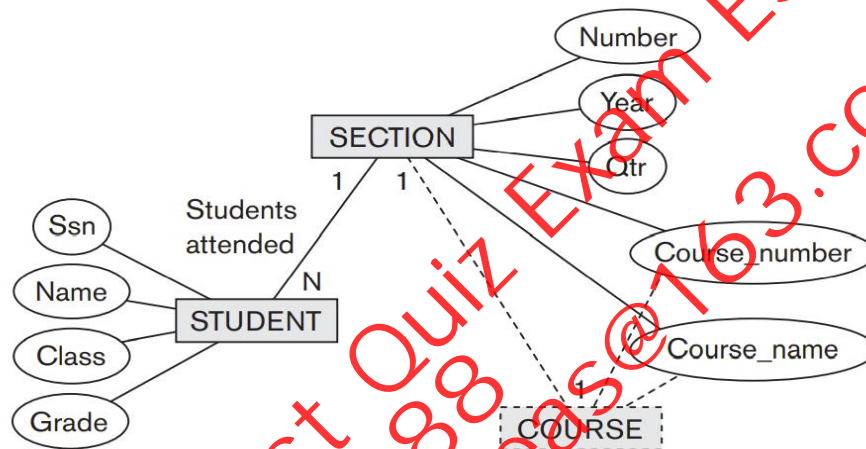
- Draw the serializability (precedence) graphs for schedule S2 based on conflicting statements in T1, T2, and T3. **(4 marks)**
- State whether the schedule is conflict serializable (conflict-free) or not. State which statement of which transaction is conflicting with other transactions. **(4 marks)**



- iii. If the schedule is serializable, write down the equivalent serial schedule. (2 mark)

3: XML Data Modeling (20 Marks)

Suppose that an application needs to extract XML documents for student, course, and grade information from a UNIVERSITY database. The data needed for this information is stored in the database attributes of the entity COURSE, SECTION, and STUDENT with One-To-One and One-To-Many relationship of SECTION with COURSE and STUDENT. Consider the following Hierarchical (tree) view for a subset of a University database and answer the following questions:



Q3(a) Create XML DTD that corresponds to the given hierarchical view. The DTD should be validated using the XML validation tool. (10 marks)

Q3(b) Create an XML instance document corresponding to the DTD you have created in the first step. The XML document must have at least one record for every element specified in the DTD. Validate the instance document against the DTD using the XML validation tool. (10 marks)

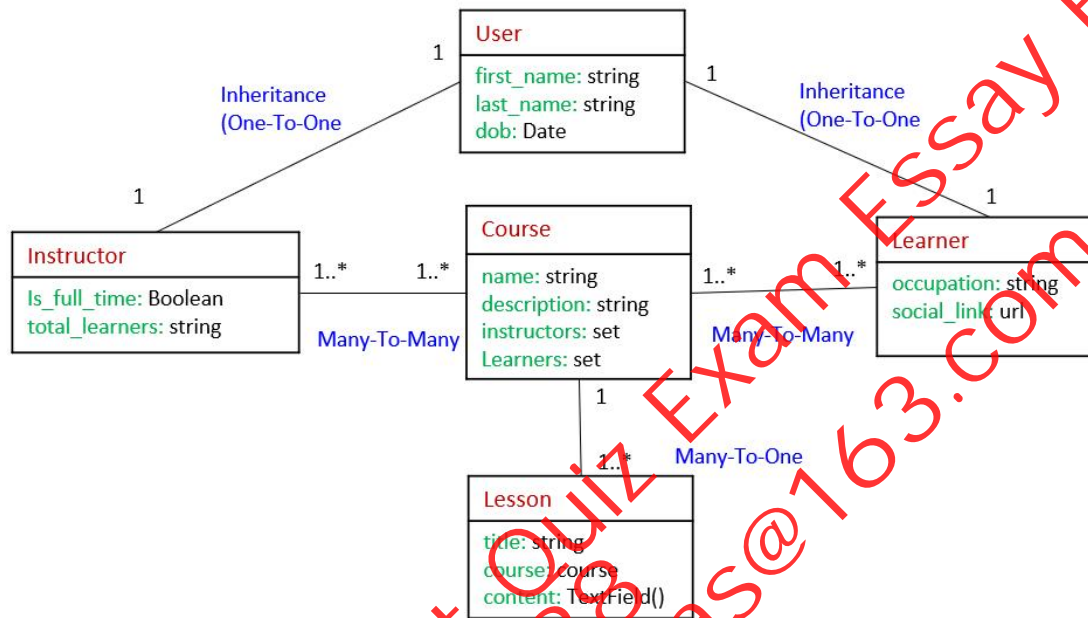
4: Object-Relational Database (20 Marks)

Suppose we want to develop a university teaching system to keep track of instructors and their teaching load. Each instructor has an office number and a unique name. Each instructor can only be either a Professor or a Teaching Assistant. A professor is identified by rank (lecturer, associate, or full professor), and a Teaching Assistant is identified by the number of years they have been assisting a course. A course is offered by a department and identified by a unique course code. An assistant can have a teaching load of at least three and at most six courses, whereas each course may have at most two assistants. A professor may teach only one course, whereas a course must have at least one and at most three professors.

Q4(a) Draw a UML class diagram for modeling information for the above system. The diagram must capture all possible attributes, primary keys, and min-max associations for each class. (10 marks)



Q4(b): Translate the following UML class diagram into a relational schema. Use the translation for subclassing most appropriate for the specified properties in the diagram. Write only the relation names and their columns; for example, User(first_name, last_name, dob); do not write the field types nor the key/foreign key constraints. **(10 marks)**



5: Data Warehousing (20 Marks)

Suppose a data warehouse for event management consists of five dimensions (Date, event, location, actors, and audiences) and two measures (total and amount_paid). Total is the count of tickets sold, and amount_paid is what the audience paid to attend an event at the given date and location. The audiences can be adults and children with different charges. The actor information includes name, age, and activities performed in an event. The date includes day, month, and year. Location information includes street, city, and theater name. Each event has a name, description, and producer information.

Q5(a) State which schema (star, snowflake or constellation) is the most appropriate to model the above data warehouse. Give your opinion of which might be more empirically useful and state the reasons behind your answer. **(4 marks)**

Q5(b) Draw schema diagram for the above data warehouse using one of the most appropriate schema diagrams. Dimension and fact tables must capture the primary keys by introducing the attribute that identifies each entity. Furthermore, a complete schema diagram must capture associations among all the entities to perform OLAP operations. **(16 marks)**



Xi'an Jiaotong-Liverpool University

西交利物浦大学

-----End of Assessment Task 2-----

Assignment Project Quiz Exam Essay Help
WeChat: cestbon_6888
Email: accoder_overseas@163.com