

**SEMESTER TWO EXAMINATION - MAY 2021
MAIN**

FACULTY:	Science, Technology and Arts
DEPARTMENT:	Computing
MODULE TITLE:	Database Systems for Software Applications
MODULE LEADER:	Konstantinos Domdouzis
SUBMISSION DEADLINE:	12 May 2021 at 09:30 GMT
SUGGESTED DURATION:	1 hour

EXAMINATION CONDUCT:

1. The [University Academic Conduct Regulation](#) outlines the behavioural expectations of candidates completing any examination.
2. Students are responsible for ensuring that they know how to submit their exam script, when the deadline is and that they submit the script in enough time before the deadline expires. It is anticipated that Blackboard will be slower around submission times.
3. It is a fundamental principle that students are assessed fairly and equitably. The [University Academic Conduct Regulation](#) defines unfair behaviour relating to an examination to be 'cheating'. The University will investigate and may sanction any acts or behaviours which breach the Code of Academic Conduct.
4. In view of the University Covid-19 response and the exceptional actions necessary to conduct off campus exams, all students are reminded that this is an individual task and that students who contact or collude with other students to complete their exam may be subject to sanction later.
5. The criteria to [Request to Repeat an Assessment Attempt](#) (RRAA) has been extended to include Covid-19 related issues - this includes illness, self-isolation, IT issues due to remote working, accessibility of learning and research materials, unexpected childcare.

INSTRUCTIONS TO CANDIDATES:

1. Whilst the release period for this exam is a 24-hour period, this exam should take you no longer than the above suggested duration to complete. Please be sensible about the amount of time you spend completing and submitting your work and make allowance for technical issues prior to the deadline.

THIS PAPER CONTAINS 5 PAGES INCLUDING THIS SHEET

2. This is an **OPEN BOOK** examination. (*you are allowed to use the Learning Materials taught during Semester 2. If you use external sources (eg. books, web-pages), you will need to reference them appropriately. Cutting-and-pasting from external sources is not allowed*).
 3. Answer **ALL** questions.
 4. Academic support will be available in two 1-hour slots between:
(11.00am-12.00pm && 16.00pm-17.00pm)
 5. It is possible that you may encounter technical issues during the exam; if you have any difficulty with IT you should consult the below student guidance document on My Hallam which contains useful information on hints and tips, contact numbers and links to support:
<https://www.shu.ac.uk/~media/home/myhallam/Guides/student-exam-guidance.docx>
 6. Any changes or clarification to the exam paper will be communicated via the module Blackboard site announcements. It is recommended that students monitor Blackboard announcements prior to submission of their final script but particularly in the first hour after release of the exam paper.
-

MODULE CODE: 55-500998

This is a blank page

MODULE CODE: 55-500998**Question 1**

- a) Describe the differences between Relational and Non-Relational Databases.

Non-relational databases are simply large stores of data all contained within one document, file, or entity (depending on what software is being used to store it).

Relational databases organise data into many different entities and describe relationships between entities' attributes which can be used to query (and edit, etc) the data from multiple linked entities.

In the early days of computing, relational databases were really the only rational choice for real-world database systems due to their ability to store data efficiently in a time where storage was expensive at high volume. However, as storage has become easier to manufacture and cheaper to purchase (or even "subscribe to"), non-relational databases have become a more viable option, and many giant companies like Google and Amazon use them today. All of their data being stored in the same place can allow for much faster querying than in relational databases, and this means for the fast-paced tech world we live in today they can be a viable database solution.

[Max no of words for 1a: 600]

[12 marks]

- b) What type of database is the MongoDB and why?

MongoDB is a non-relational database. The reason for this is that it is aimed at users who need an extremely flexible and easily expandable database which can be queried quickly. (<https://www.mongodb.com/non-relational-database>)

[Max no of words for 1b: 100]

[2 marks]

- c) What is an associative array and in which type of databases is it used?

An associative array is a type of array, similar to a hash map, which holds pairs of keys and values (for example a house number and the relevant house owner name). They are used in both relational and non-relational databases due to their low-resource-cost ability to link some pieces of data together.

[Max no of words for 1c: 200]

[3 marks]

- d) NoSQL databases have less rigid or even nonexistent data model restrictions. Do you think this characteristic of NoSQL databases makes them appropriate for use in complex scenarios that involve large amounts of data? Please explain your answer.

NoSQL databases' lack of data model restrictions, or rigid schema, allows them to be more flexible than traditional SQL databases which I think is an advantage in complex scenarios. This is because the structure of data being entered into the database can be quickly changed whenever needed to suit the needs of the rest of the project, and the database will still function the same. Another benefit of NoSQL database in this kind of situation is that large amounts of data will not make the database slow to use as due to its simplicity it will remain fast to query.

[Max no of words for 1d: 300]

[3 marks]

Question 2

- a) In a Distributed Database Management System (DDBMS), which rule of Fragmentation enables the minimisation of data redundancy? Explain the specific rule.

The disjointedness rule of fragmentation ensures that after the process of fragmentation, no record from the original table becomes a member of more than one fragment – thus ensuring there is no duplicate (redundant) data. The notational expression of this rule shows that the intersecting set of all fragments should not contain anything, as there is no record present in more than one fragment.

(<https://www.exploredatabase.com/2017/05/correctness-rules-for-verifying-fragmentation-in-distributed-database.html>)

[Max no of words for 2a: 150]

[2 marks]

- b) Explain the way the replicated data in a Distributed Database Management System (DDBMS) achieve data consistency.

In a database system, the ways that data can be added and changed are restricted. This, alongside relationships between attributes of tables ensures that no data can be changed in such a way that any other data loses its meaning or validity; ensuring the structure of the data stays consistent, well-organised and queryable.

[Max no of words for 2b: 150]

[2 marks]

- c) Describe the type of Distributed Database Management Systems (DDBMSs) that enable scalability. How is concurrency achieved in a DDBMS?

Concurrency is achieved in a DDBMS using strict data schemas and locking of data based on time and activity to try and ensure data is not edited from multiple places at once. An example of this is timestamp currency control algorithms, which use the timestamp of a transaction (i.e. a write to the database where data was inserted or updated) to ensure that it is applied to the data in the correct order, i.e. before any transactions with later timestamps. This may not prevent data being unnecessarily updated or created multiple times, but should ensure, given a sufficient storage of transaction history, the ability to roll back changes if necessary.

[Max no of words for 2c: 400]

[6 marks]

- d) The concurrent execution of various transactions can cause problems, called anomalies. One of these anomalies is the Inconsistent Read. Describe this anomaly.

An inconsistent read can occur in a database system when data is being queried (read) from one location or endpoint in the system, and concurrently edited (written) from another point in the system. If timed correctly (or, incorrectly) it is possible that the query fetches some of the data pre-edit and the rest post-edit, meaning the record it displays may be completely invalid, composed of out-of-date data.

[Max no of words for 2d: 150]

[4 marks]

- e) A Database Management System (DBMS) must serve many applications and serve requests from many different users. These operations are expressed in the form of transactions. Describe the mechanisms used by a DBMS in order to achieve auditing during transaction processing.

There are six main ways that DBMS use to achieve auditing of databases:

1. **DBMS traces** allow database admins to start tracking specific types of activities, e.g. all updates to the data.
2. **Temporal capabilities** are offered by some modern DBMS and allow for modification of data to be audited by means of a dedicated history table related to the relevant table in the real database.
3. **Transaction log** files are how they sound – most DBMS automatically log every transaction in files that are accessible later if needed for an audit of who did what, when.
4. **Network auditing** is a process that listens for SQL being sent over a network and logs it.
5. **Hand-coded audit trails** add dedicated columns in the database that track data such as the user and date that last modified that record.
6. **Direct server auditing** captures all SQL requests sent to the server for viewing, without a potentially storage-expensive log of all transactions or history of tables.

(<https://www.dbta.com/Editorial/Think-About-It/Improving-IT-Security-With--Database-Auditing-Techniques-112755.aspx>)

[*Max no of words for 2e: 300*]

[6 marks]

END