

MANDATORY ASSIGNMENT #2

IN-3020/4020

Due Friday 8th May 2020

(by midnight)

The exercises in this mandatory assignment must be solved and submitted individually.

Completion and submission of the mandatory assignment must take place in accordance with current guidelines of the Department of Informatics (Institutt for Informatikk), see:

www.uio.no/studier/admin/obligatoriske-aktiviteter/mn-ifi-oblig.html

Submission of the mandatory assignment will be taken as the confirmation that guidelines have been read and understood.

Deadline: Friday, May 8th at 23:59 (by midnight). Use **Devilry** and contact us if you experience problems delivering!

The deadline is absolute, and deferral will only be granted in accordance with the mandatory assignment rules (obligreglement).

All questions must be answered in order to get an approval.

IMPORTANT NOTE: The questions are formed in a manner that will give you an idea about the take home written exam.

Exercise 1 – Transaction Management, ACID characteristics

1.1 Concurrency & ACID characteristics

We know that concurrency can threaten certain ACID characteristics. Explain which ACID characteristics are endangered due to concurrency. Note that you **must explain why** these characteristics are challenged!

1.2 Concurrency example

- a. We are going to be looking at an example of possible conflicts between **2 concurrent transactions** in a vacation reservation system.

A customer (guest) can log in to the vacation reservation system and choose one flight that is marked as available (data element F) and one hotel room marked as available (data element H) from their respective databases. When the customer has decided & booked, both entities will be marked unavailable.

Detail two such concurrent transactions T_1 and T_2 by dividing each into their read/write operations. Note that both transactions must have all necessary read and write operations.

Ignore the date or period. Assume that we are operating with two data elements (flights and hotels marked as available/not-available) only.

- b. Show at least two possible interactions between the transactions with different interleaving of their operations where at least one of them demonstrates a correct and at least one demonstrates an incorrect result. Note that you should say something about what **your correctness (integrity) criterion** is.
- c. Explain what the two alternative runs imply of possible solutions to the concurrency problem. Reason for at least TWO possible solutions.

1.3 Logs

We have many approaches and mechanisms addressing the challenges of concurrency. Logging is one of them.

- a. List up and explain what the three types of logging mechanisms are and compare them with respect to what they do and what their consequences are.
- b. Which ACID characteristics does logging solve? Explain why, giving your reasoning.
- c. Why do we have checkpointing? What problem does it solve? How is it different from the three logging alternatives?

Exercise 2 – Concurrency Control, Serialization

2.1 Execution plans

- a. Take your example in assignment 1.2, remember that it is NOT going to be from the slides and lectures, and create 3 different example execution plans.
- b. Explain which are serial or serializable and why.

2.2 Conflict serializability, locks

- a. What kind of conflicts can one have in an execution plan? List them up and explain in your own words why (if necessary, with examples).
- b. Draw the precedence graphs of your 3 execution plans in 2.1 (a).
- c. Just by looking at the precedence graphs: Which one of them are conflict serializable?
- d. Explain the reasoning behind 2 phase locking. Note that you are not being asked for the formulation of 2PL only but for the reasoning behind the mechanism. What does it help us with? Why?

Exercise 3 – Isolation levels

3.1 Snapshot Isolation and related mechanisms

What is Snapshot Isolation, what is First Update Wins, which challenges they address, and how do they compare with or relate to time-stamping and versioning mechanisms? Discuss, explaining in your own words (if necessary, use examples/sketches).

3.2 Isolation levels and concurrency mechanisms in practice

This is a “look up, learn and share” exercise: What concurrency and isolation mechanisms do Postgres, MySQL and Oracle support? Look them up, list them and compare!

Note: Use the most recent versions.

Beware: There will be several **types** of DBMS some of them offer that focus on different capabilities/mechanisms.

Exercise 4 – NoSQL DBMS

The **Books** table is given at the end of the question. Take a look at it.

4.1 Is the Books table a relational table? If not, explain why.

4.2 Model the data in the given table using JSON. Explain your modelling choices (especially related to normalization/denormalization of data based on the type of relationships).

The following assumptions are given:

- Authors are registered with their full names (e.g., “Joe Doe”)
- Publishers are recorded with name (e.g., P1), contact email (e.g., contact@p1.no), and phone (e.g., 565489)
- Ratings can be of different types, quantitative/numerical (stars on a range from 1-5) or qualitative (comments such as “good”, “excellent”, etc.)

4.3 Assuming all the documents produced under item 4.2 are in the same collection, write the following basic queries in MongoDB:

- a. Retrieve all documents in the collection
- b. Retrieve the documents containing only books that are available (“Y”)
- c. Retrieve the documents containing only books for which there is information about their availability (“Y” or “N”)
- d. Retrieve the documents containing only books that appeared before 2019 for which there is information that they are not available (“N”)
- e. Retrieve the documents containing books whose ratings consisting of a 5 star.
- f. Retrieve the documents containing books published by publisher “P1” (Do this query only if your data is denormalized)

References:

- Use an online JSON validator to validate the documents (e.g., <https://jsonformatter.curiousconcept.com/>) before including them in MongoDB
- Data model design in MongoDB: <https://docs.mongodb.com/manual/core/data-model-design/>
- MongoDB query documents: <https://docs.mongodb.com/manual/tutorial/query-documents/>

The Books table:

Title	Authors		Publishers	Available	Ratings	Sold copies
Into the dark	Joe Doe Ann Smith Bill G.	2019	P1, contact@p1.no, 565489 P2, email@p2.no, 4654575	Y	5, 3, good, 2.5	80
Start again	Ann Smith John A.	2019	P3, john@p3.no, 646548	Y	excellent, 4.5, 2	60
Try hard	John A.	2018	P2, email@p2.no, 4654575	N		10
Job done	Per Franck	2016	P1, contact@p1.no, 565489 P3, john@p3.no, 646548	N	4, 5	70
Into the dark	Bill G.	2020	P2, email@p2.no, 4654575			
Not again	John A.			N		

>END OF MANDATORY ASSIGNMENT #2<
and GOOD LUCK!