Course guide
340457 - DABD-I7P23 - Databases Design and Administration

Last modified: 17/05/2023

Unit in charge: Vilanova i la Geltrú School of Engineering
Teaching unit: 723 - CS - Department of Computer Science.

Degree: BACHELOR'S DEGREE IN INFORMATICS ENGINEERING (Syllabus 2018). (Compulsory subject).

Academic year: 2023 **ECTS Credits:** 6.0 **Languages:** Catalan

LECTURER

Coordinating lecturer: Jordi Esteve

Others: Jordi Esteve

PRIOR SKILLS

The ones of the Programming and Specification subjects.

REQUIREMENTS

Have passed INEP and PROP or at least have the knowledge taught in INEP and PROP.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
1. CEC01. Ability to have a thorough understanding of the fundamental principles and models of computation, ability to apply the principles to interpret, select, evaluate, model, and create new concepts, theories, applications and advance the technological development related to computing.
2. CECO5. Ability to acquire, obtain, formalize and represent human knowledge in a computable form for problem solving by using a computer system in any scope, particularly those related to aspects of computation, perception and action in intelligent environments.
3. CEFC8. Ability to analyze, to design, to construct and to maintain applications in a well built, secure and efficient way choosing the most adequated paradigms and languages.
4. CEIS4. Ability to identify and analyze problems and design, develop, deploy, test and document software solutions based on an adequate knowledge of theories, models and techniques.
5. CESI2. Ability to determine the requirements of information and communication systems of an organization paying attention to safety aspects according to security and compliance with regulations and legislation.

TEACHING METHODOLOGY

Mix sessions with theoretical developments and practical exercises.

LEARNING OBJECTIVES OF THE SUBJECT

Ability to design relational databases for specific applications. Ability to effectively manage relational databases, mastering the advantages and disadvantages of the main available mechanisms. Ability to design and manage non-relational databases, with mastery of advantages and disadvantages compared to relational systems.

STUDY LOAD

Type	Hours	Percentage
Hours small group	30,0	20.00
Self study	90,0	60.00
Hours large group	30,0	20.00

Total learning time: 150 h

CONTENTS

1. Relational algebra and SQL: Selection, projection, composition (join), grouping, materialization, renaming, sorting

Description:
Sets and relationships between sets. Relations, tuples and attributes. Selection, projection, composition (join), grouping, materialization, renaming, sorting. NULL value.

Full-or-part-time: 14h
Theory classes: 2h
Laboratory classes: 6h
Self study : 6h

2. Conceptual modeling: functional dependencies, keys, normal forms, and normalization

Description:
Analogies and differences between UML class diagrams and relational diagrams. Transformation UML diagram to relational schema. Functional dependencies, keys, normal forms (1NF, 2NF, 3NF and Boyce-Codd NF) and normalization. Over-normalization and de-normalization.

Full-or-part-time: 30h
Theory classes: 12h
Laboratory classes: 8h
Self study : 10h

3. Indices: Hashing and B+ trees

Description:
B, B+ and "hash" indexes. Clustered index. Partial index. Covering index. When not to create indexes.

Full-or-part-time: 8h
Theory classes: 2h
Laboratory classes: 2h
Self study : 4h

4. Transactions: atomicity, coherence, isolation and durability (ACID). Triggers

Description:
Persistence or durability: disk allocation considerations (mirrors, logs), crash recovery. Transactional process: atomicity. Standard levels of isolation: advantages and disadvantages of each one; how to choose the most suitable one. Triggers

Full-or-part-time: 13h
Theory classes: 5h
Laboratory classes: 4h
Self study : 4h

5. Process of an SQL query and optimization

Description:
Implementation of basic relational operations; their combinations. Generation of plans. Different Access Paths available. Different ways of doing joins (nested loops, sort-merge, hash join). Predictions and statistics.

Full-or-part-time: 8h
Theory classes: 2h
Laboratory classes: 2h
Self study : 4h

6. How to manage a database

Description:
Database Administrator Principles: Think globally; act locally. Agree to negotiate. Trust the specialist. It costs more to start than to continue. Spread the work.

Full-or-part-time: 2h
Theory classes: 1h
Self study : 1h

7. Non-relational databases

Description:
SQL and relational systems: advantages and disadvantages. Scaling UP/Scaling OUT. CAP principle. BASE properties. Recent "no-SQL" and "not-only-SQL" alternatives: relevant considerations.

Full-or-part-time: 10h
Theory classes: 2h
Laboratory classes: 4h
Self study : 4h

Partial and final tests

Description:
Partial (2h) and final (3h) tests

Full-or-part-time: 5h
Guided activities: 5h

Project

Description:
Individual project where an application that manages a relational database will be created to solve a data management problem. The following tasks must be performed:
- Description of the problem with a glossary of the terms that appear
- UML diagram that describes the problem
- Obtain the relational schema from the UML diagram
- Optimization of the DBMS used according to the characteristics of the server and the estimated simultaneous connections
- Script used to add millions of real data in the DB and additional indexes created to optimize queries (if they are necessary)
- Decide what technology the application will use
- Application that manages the entire relational schema implemented (or a part if the relational schema is very large)

Full-or-part-time: 60h
Guided activities: 60h

GRADING SYSTEM

C1 = First Test about DB Design. Individual written test (2 hours).
Second Test. Individual written test (3 hours) which integrates knowledge and skills of the entire course:
C2a = 1st part about DB Desgin (optional)
C2b = 2nd part about DB Administration
Act1 = Activity1 + lab exercices 1st part
Act2 = Activity2 + lab exercices 2nd part
Pra = Grade obtained from the individual practice (an application using a RDBMS).

Final Grade = 0,2*max(C1, C2a) + 0,2*C2b + 0,2*Act1 + 0,1*Act2 + 0,3*Pra

The presentation of the practice will be mandatory to pass the course; otherwise, the final grade of the whole subject will be 'NP'.

The Review Test is a written test of maximum 3 hours and made up of two parts, C3a and C3b, being able to do both or only one, which replace grades C2a and C2b.

EXAMINATION RULES.

The written tests (Control 1, 2 and 3), the lab exercises, the activities and the project are individual.

In each Laboratory session, an attendance control will be done.

The date for the delivery of an exercise/activity/project will always be a deadline.