Advanced Databases

(CIS 4930 / CIS 6930)

Dr. Markus Schneider

Department of Computer & Information Science & Engineering (CISE)

Course Syllabus[1] – Fall 2016

November 15, 2016

Contents

---

[1]Note that this syllabus has a "dynamic" character. That is, it can (and will be) updated and extended at any time without prior notice. Hence, from time to time you should check the syllabus.

1   General Information

Credits: 3

Prerequisites: General knowledge of data structures and algorithms; general database
                      knowledge (for example, obtained in classes such as CIS 4301 or COP 5725);
                      web programming skills (e.g., PHP)

Instructor: Dr. Markus Schneider

Lecture times: Mondays, Wednesdays, Fridays: 12:50 pm - 1:40 pm (6th period)

Location: CSE 119

Office hours: Mondays, Wednesdays, Fridays: 9:45 am - 10:45 am
                      or by appointment

Office: CSE Building, room E450

Course web site: http://www.cise.ufl.edu/~mschneid/
                      Teaching/CIS4930+CIS6930_Fall2016/CIS4930+CIS6930_Fall 2016.html

Phone: (352) 505-1584

Email: *mschneid@cise.ufl.edu*

2   Course Objectives

Relational database systems are (still) the most widespread data management and processing
technology used in industry and public services. They usually deal with simply structured data and
administrative, business, and financial applications. But new kinds of emerging data with their
different nature and often complex internal structure as well as new emerging applications based
on these complex data have revealed that standard relational database systems are not the
appropriate tools to handle them and that instead *novel data models* and *non-standard*, *advanced
database systems* are required to adequately support them.

This lecture provides a closer insight into a selected collection of advanced database systems and
investigates their underlying data models, query languages, concepts, principles, techniques,
functionalities, architectures, and special features, and compares them to established relational
database systems. Examples of such advanced systems are PostGIS, MongoDB, and SciDB. In
addition, the goal is to design and implement a meaningful real-world application based on each of
these systems in order to demonstrate their usefulness and power. Additionally, students learn to
perform a literature study, write a technical document, and design and hold a PowerPoint
presentation. Finally, students learn to work in and be responsible for a group since group work
will be the basis of all deliverables. This will increase their social competence as well as their
communication and discussion skills.

Each group will consist of four students. The students in class may form the groups themselves. In
more detail, this class includes the following teaching and learning objectives:

- *Assignment, installation, and demonstration of an advanced database system*. An
  advanced database system from a pre-defined list is assigned to each group. Group
  priorities will be taken into account. Each group has to install the assigned advanced
  database system on a laptop or a PC. Note that laptops or PCs cannot be provided by the

university or the CISE department for this task so that laptops or PCs have to be taken from group members. The objective is to learn how to install a software system on a computer, cope with installation problems, and also provide installation requirements such as particular operating systems on which the advanced database systems run. Finally, each group will provide an in-class demonstration of their system.

- "*Students teach students*". Only at the beginning the instructor will hold lectures. Afterwards, the student groups will take over. The objective is that each group will learn how to present their assigned system to all the other students in class by means of a PowerPoint presentation. In this manner, students will teach students as well as practice and improve their presentation and communication skills.

- *Performing a literature study*. Before one can deal with a new topic or system, one has to become aware of the state of the art regarding this topic or system. Therefore, an objective is that students learn how to perform a literature study and process the found references and papers. The literature of interest can relate directly to the topic or system itself but also be centered around the topic or system. For example, if a system rests on a particular data model, scientific literature about this data model would be very helpful. References will be kept in BibTeX format.

- *Design and implementation of an application that showcases the special features of the assigned system.* Each group will develop and implement a meaningful web-based application that demonstrates the particular features of the assigned advanced database system. The objective is to learn to perform software development and implementation on top of a new database software. Each group will decide about the deployed programming languages and software packages.

- *Writing a LaTeX document about the assigned system*. At the end of the semester, each group will provide a well written and well formatted document that describes the main features of the assigned advanced database system as well as the designed and implemented application. LaTeX is the only word processing tool that is allowed for writing this document.

More details will be provided in class. By providing a balanced view of theory and practice, the material covered should allow the student to understand and use advanced database systems.

3   Course Information

3.1   Textbooks and Readings

No textbooks and readings are required. We will read original technical and scientific literature.

3.2   Staff

Instructor:

Dr. Markus Schneider
Office: CSE Building, Room E450
Tel.: (352) 505-1584

E-mail: `mschneid@cise.ufl.edu`

Teaching Assistants:

Please take the names, the offices, the telephone numbers, the office hours, and the email addresses of the teaching assistants from the course website.

4 Course Assessment

4.1 Project Deliverables and Grading

There will be *no* exams and *no* homework assignments. Student group assessment will be based on the following six deliverables (in the following *ADBS* means *Advanced Database System*):

1. Literature study about the group-assigned ADBS described in a text document written in LaTeX and with references in BibTeX format and papers in PDF format (5 %)

2. ADBS installation and demonstration (10 %)

3. PowerPoint presentation about the group-assigned ADBS in front of the class (25 %)

4. Application design showcasing the outstanding features of the group-assigned ADBS and described in a text document written in LaTeX with references in BibTeX format (15 %)

5. Application implementation and demonstration (25 %)

6. Detailed overview text document about the group-assigned ADBS written in LaTeX with references in BibTeX format (20 %)

The grading criteria of the literature study (point 1) are:

1. Does the literature study provide the start of the art?

2. Does the document define literature categories?

3. Is the literature classified according to these categories?

4. Have the strengths and especially weaknesses of each reference been described?

5. Has a connection or comparison been made between the reference and your system?

The grading criteria of the ADBS system demonstration (point 2) are:

1. Organization of the system demonstration (35%)

   - Was the demonstration well structured?

   - Did the demonstration have a clearly defined goal/focus/message?

   - Did the demonstration give the essential system features and facts?

   - Was the demonstration "complete", or were there any important aspects of the system that were missing or only partially covered?

2. Clarity and comprehensibility of the talk (35%)

   - Was the demonstration easy to follow?

- Did the presenters express their thoughts with a clear, loud, and expressive voice?
- Was the demonstration held in a fluent and comprehensible manner?
- Were the transitions between the group members smooth?

3. Knowledge and depth of the presented material (30%)

- Was the material covered at a depth that is adequate for the class?
- Did the presenters give the impression that they are knowledgeable and know what they are talking about?
- Did the presenters provide the most important information and facts?

The PowerPoint presentation (point 3) should answer the following questions:

1. What are the characteristic features of your system?
2. What are the applications your system aims at?
3. How does your system aim at these applications and their requirements?
4. What is the underlying (formal?) data model?
5. What are the operations supported by this data model and this system and how do they work?
6. What kind of query languages are provided for your system and how do they work?
7. How can application programs be written on top of your system?
8. Which are the most important implementation concepts of your system and how do they work?

The grading criteria of the PowerPoint presentation are:

1. Organization of the slides (35%)

- Were the presentation slides well structured?
- Did the presentation slides have a clearly defined goal/focus/message?
- Did the presentation slides give the essential facts and results?
- Was the presentation "complete", or were there any important aspects of the topic that were missing or only partially covered?

2. Clarity and comprehensibility of the talk (35%)

- Was the presentation easy to follow?
- Did the presenters express their thoughts with a clear, loud, and expressive voice?
- Was the presentation held in a fluent and comprehensible manner?
- Were the transitions between the group members smooth?
- Did each group member speak for the same amount of time?

3. Knowledge and depth of the presented material (30%)

- Was the material covered at a depth that is adequate for the class?
- Did the presenters give the impression that they are knowledgeable and know what they are talking about?

- Were the presenters knowledgeable about the related literature?
- Did the presenters provide the most important information and facts?

The grading criteria of the application specification (point 4) are:

1. Introduction and Motivation (40%)
   - How well are the advanced database system (ADBS) and the intended application introduced?
   - How good is the motivation that the intended application fits to the ADBS? Is it realistic?

2. Application specification (50%)
   - How well are the details of the application worked out?
   - How well is the database structure described and explained?
   - How well is the web-based user interface designed and worked out?
   - How good is the description of the interplay between the ADBS and the application?

3. System and implementation environment (10%)
   - How well is the system configuration (single laptop or PC, multiple laptops or PCs) described?
   - How well are the implementation resources (ADBS, operating system, programming languages, script languages, tools, public domain software packages, etc.) described?
   - How well is the interplay between all the used system and implementation components described?
   - What are the data sources that will be taken into account? How large are they?

The grading criteria for the detailed overview text document (point 6) are:

1. Introduction (10%)
   - How well is the ADBS introduced from a broader perspective?
   - How well are important applications sketched from a broader perspective that would benefit from using the ADBS?
   - How well is the need for the ADBS motivated?
   - How well are the main features of the ADBS sketched?
   - How well is the layout, formatting, spelling, and grammar of this chapter?

2. Related Work (15%)
   - How well is the literature described that refers to the theoretical foundations underlying the ADBS?
   - How well is the literature described that refers to the ADBS itself?
   - How well is a comparison made to competitors?
   - How well is the layout, formatting, spelling, and grammar of this chapter?

3. System Description (30%)

- How good is the overview of the ADBS presented and motivated?
- How well is the architecture of the system described and motivated?
- How well is the underlying data model described and motivated?
- How well are the available operations on the data described and motivated?
- How well is the query language of the ADBS described?
- How well are all system aspects illustrated by examples and figures?
- How well are the characteristic features of the ADBS described?
- How well is the layout, formatting, spelling, and grammar of this chapter?

4. Application Description (40%)

- Has the application chosen been well described and motivated?
- Has there been an argumentation why the application chosen showcases the features of the ADBS?
- How well have the details of the application been worked out with respect to its design and implementation?
- How well is the database structure described and explained?
- How well is the web-based user interface designed and worked out?
- How good is the description of the interplay between the ADBS and the application?
- How well is the layout, formatting, spelling, and grammar of this chapter?

5. Conclusions (5%)

- Has a summary of the previous chapters been given?
- What have you learned from dealing with the ADBS?
- What is your personal assessment of the ADBS as a group?
- How well is the layout, formatting, spelling, and grammar of this chapter?

During the semester the student can only earn *points* and *not* grades. At the end of the semester the weighted sum of all points is mapped to a grade. The maximum number of points that can be achieved for any deliverable is always 100 points.

The student's performance *p* will be calculated according to the following formula:

$$p = 5 \cdot \frac{ls}{100} + 10 \cdot \frac{id}{100} + 25 \cdot \frac{pp}{100} + 15 \cdot \frac{ad}{100} + 25 \cdot \frac{ai}{100} + 20 \cdot \frac{td}{100}$$

where the variables have the following meaning:

- p student's performance in percent
- ls received number of points for literature study

- *id*   received number of points for ADBS installation and demonstration
- *pp*   received number of points for PowerPoint presentation
- *ad*   received number of points for application design
- *ai*   received number of points for application implementation
- *td*   received number of points for overview text document

Point assessments (gradings) will be changed only when an assessment error or an error in the grade spreadsheet has been made; *negotiation is not allowed*. If a student group thinks an error has been made, the group should let the instructor or the TAs know about their doubt in person or by email. The entire deliverable will then be reevaluated. *A student group must submit a deliverable for reevaluation within 5 workdays from when the evaluation of that item was returned in class or posted in the grade spreadsheet*. Based on the group's performance, the following grading policy will be applied to all group members:

Student's Performance p	Letter Grade
>94-100	A
>88-94	A−
>82-88	B+
>76-82	B
>70-76	B−
>64-70	C+
>58-64	C
>52-58	C−
>46-52	D+
>40-46	D
>34-40	D−
0-34	E

Note that the instructor takes the right to assign point assessments to individual group members that deviate from the point assessment for the group. This will be the case if it turns out with respect to any deliverable that a group member has performed much better or much worse or contributed much more or much less than the other group members. In particular, this measure is supposed to prevent that individual group members "hide" within their group and let the other group members work for them.

Based on the graded deliverables by then, students will regularly obtain information about their individual grades, current performance, and overall grade in a class grade spreadsheet if student anonymity is guaranteed.

A C− will not be a qualifying grade for critical tracking courses. In order to graduate, students must have an overall GPA and an upper-division GPA of 2.0 or better (C or better). Note: a C−

average is equivalent to a GPA of 1.67, and therefore, it does not satisfy this graduation requirement. For more information on grades and grading policies, please visit the web page: https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx.

4.2   Class Participation

The instructor highly recommends a regular class attendance. Much relevant information will be provided in class only. The instructor would also like to point out that he highly advocates punctuality for the lectures.

4.3   Late Policy

Be aware that there is no late policy for this class.

5   Other Important Issues

5.1   Academic Honesty

Students are required to respect the ethical standards for academic honesty established by the Office for Student Judicial Affairs. The University of Florida student body voted in Fall 1995 to approve the following *Student Honor Code*:

> *We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honesty and integrity.*

Work submitted must be produced individually by each student, except for tasks explicitly assigned to a group by the instructor. All work submitted individually in the form of exams, homework, presentations, reports, software projects, etc., is subject to the following implicitly or explicitly **required** pledge:

> *On my honor, I have neither given nor received unauthorized aid in doing this assignment.*

Violations of academic honesty and integrity in this course will not be tolerated. The instructor will deal strictly with any violations. Since ethical behavior in science and engineering is equal in importance to specific knowledge, the instructor will assign a non-passing letter grade to students who violate academic honesty standards, regardless of the violator's grade performance in exams, homework, quizzes, and other assignments. Official sanctions issued by the Office of Student Judicial Affairs **will become permanently noted in the student's official transcript**.

The instructor's advice to the student is: immerse yourself in the class, learn the material, do your tasks (homework, presentation, report, implementation, etc.) yourself. The benefit and enjoyment you will receive as a result of hard work will be much more valuable than any penalty you might receive as a result of cheating.

5.2   Class Rules

For this class several rules hold which should be observed by the student:
1. If a student should have a problem that could have a negative influence on the student's class performance like sickness or project group conflicts, the student should talk to the instructor on

time before it is too late in order to find a solution.

2. The instructor attaches great importance to punctuality. Hence, each student should come in time to the class.

3. Class-relevant announcements are made also in the class. If a student is unable to attend a class, s/he should ask a fellow student or pass the instructor's office during the office hours regarding announcements in the previous class.

5.3   Students with Disabilities

Students requesting classroom accommodation must first register with the Disability Resource Center (DRC) at the beginning of the semester. The DRC will provide documentation to the student who must then provide this documentation to the instructor when requesting accommodation.

5.4   Where to Get Class-Related Information

The latest class-related information is given at the beginning of each class when announcements are made. Some information related to this class will be provided on the course web site. This especially relates to the slides of the different lectures. The web address of the home page of this course is listed at the beginning of this syllabus.

5.5   Final Advice

The students who get the most out of this class will be the ones who put in the most effort. If you want to do well, come to all the lectures, work hard, be actively involved, and contribute to your group's success. If you are having difficulty, you owe it to yourself to get help. The instructor and the teaching assistants sincerely want all of you to do well. The instructor will try his best to make the course as interesting and stimulating as possible and an enriching experience for you.

Markus Schneider