Comment: This course is the merger of two previously independent courses,

- Knowledge Representation and Ontologies (Diego Calvanese)
- Information Integration (Werner Nutt)

The courses have some overlap, but differ in the learning outcomes, the approach to teaching, and the assessment.

Over the summer, we will redesign the course to make it a coherent single course. Due to time constraints, this was not yet possible.

COURSE NAME: Ontology and Database Systems

COURSE CODE

LECTURER: Diego Calvanese (8CFU) and Werner Nutt (4 CFU)

TEACHING ASSISTANT: TBD

TEACHING LANGUAGE: English

CREDIT POINTS: 12

LECTURE HOURS: 72

EXERCISE HOURS: 36

OFFICE HOURS LECTURER: Diego Calvanese: during the lecture time span Friday 15:00–17:00, Faculty of Cs, POS Building, piazza Domenicani 3, office 2.07; outside of the lecture time span by previous email appointment.

Werner Nutt: TBD

OFFICE HOURS TEACHING ASSISTANT: TBD

PREREQUISITES:

- Notions of first-order logic as taught in an introductory BSc course on Mathematical Logic;
- Notions of relational databases as taught in an introductory BSc course;
- Notions of Java programming and SQL with JDBC Connectivity.

Students should be familiar with relational databases, and, on at least an introductory level, with logic and elementary algorithms. In particular, knowledge of SQL and relational algebra and of the fundamental concepts of the theory of NP-completeness are required. At FUB, this material is taught in the courses Introduction to Databases and Theory of Computation.

OBJECTIVES

The aim of the course is to provide students with an understanding of the formal foundations of classical logic-based knowledge representation languages,
with an overview of the reasoning methods for them, and of the application of techniques developed in knowledge representation to classical data management problems. Most of the course will focus on description Logics and on ontology languages.

The course introduces the formal concepts from the area of databases by which information integration problems are modeled. It presents techniques for evaluating queries in this setting, mapping schemas to each other and assessing the quality of query answers.

In addition to studying the technical material, students will train fundamental mathematical skills such as giving formal definitions, formulating theorems, and proving or disproving formal statements.

SYLLABUS
Modeling information through ontologies, Description Logics, Query answering over databases and ontologies, Ontology based data access, Reasoning in the DL-Lite family, Reasoning in the ALC family

- Basics of Relational Database Theory
- Modeling Incomplete Information
- Modeling Information Sources: Global as View, Local as View
- Query Semantics and Query Planning
- Sources with Access Limitations (Forms, Web Services)
- Data Exchange
- Schema Mapping

TEACHING FORMAT
The course is organized as frontal lectures on the course topics, possibly complemented by monographic seminars that serve as a starting point for discussing the techniques involved. During lab sessions the students will familiarize with the usage and internals of state-of-the-art tools for managing and querying relational data sources through an ontology, and will work on a project.

Lectures and exercises

ASSESSMENT
The exam consists of:
- a project [30 % of mark]
- a final oral or written exam [70 % of mark]

Both parts have to be passed to pass the exam, but they can be taken independently of each other.
In case of a positive mark, the part that has been passed will count for all 3 regular exam sessions of the Academic Year (i.e., if the student fails or does not take, e.g., the oral exam, (s)he keeps the project and only needs to retake the oral exam).

The final mark will be based on exercises and on a written exam. Students will submit solutions for exercises and present their solutions in class.
Students who do not submit exercises will be assessed on the exam alone. For students who do submit exercises, the final mark will be based on both the exam mark and the exercise mark. In this case, a weighted average of the exam mark (70%) and the exercise mark (30%) will be computed. If this average is greater than the exam mark, the average will be the final mark. Otherwise, the exam mark will be the final mark. The exercises will be taken into account, independently of when the student takes the exam.

**READING LIST**


Suggested Readings:

- Foundations of Databases: S. Abiteboul, R. Hull, and V. Vianu, Addison Wesley, 1995 (available online)
- Principles of Data Integration: AnHai Doan, Alon Halevy, Zachary Ives, Morgan Kaufmann, 2012
- Research articles distributed in class and on the course webpage

Lecture notes and additional reading material covering the course topics will be provided during the course and made available in the course web page.

**SOFTWARE USED**

- Protégé ontology editor
- MySQL or Postgres database engine
- Ontology-based Data Access Tools from [http://ontop.inf.unibz.it/](http://ontop.inf.unibz.it/)

**LEARNING OUTCOME**

Students will acquire an understanding of the advanced languages, methodologies, and the use of knowledge representation techniques, also in the context of accessing and querying information sources. Automated reasoning techniques and formal semantics will be understood for these languages.

After the course, students will

- be familiar with the theoretical concepts underlying databases and query languages
- know how to model information integration scenarios
- know how to interpret queries in an information integration setting and develop solutions that compute query answers
- be able to formulate theoretical statements about database queries and to prove/disprove them.
http://www.inf.unibz.it/~calvanese/teaching/odba/

http://www.inf.unibz.it/~nutt/InfInt1112/