جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Course Syllabus: Applied Ontology - CS 322

Division	Computer, Electrical and Mathematical Sciences & Engineering
Course Number	CS 322
Course Title	Applied Ontology
Academic Semester	Fall
Academic Year	2018/2019
Semester Start Date	08/26/2018
Semester End Date	12/11/2018
Class Schedule (Days & Time)	04:00 PM - 05:30 PM | Sun Wed

Instructor(s)

Name	Email	Phone	Office Location	Office Hours
Robert Hoehndorf	robert.hoehndorf@kaust.edu.sa	+966128081643	4222, 3, Ibn Sina (bldg. 3)	by appointment only

Teaching Assistant(s)

Name	Email

Course Information

Comprehensive Course Description	The course covers advanced topics in conceptual modelling, data management, integration and analysis, all of which have applications in data-intensive disciplines such as biology, biomedicine and others. The aims of the course is to provide an in-depth understanding of the state of the art in formal ontologies, including their role in integrating and analyzing data. While Knowledge Representation and Reasoning (CS213) introduced basic logic formalisms that can be used to express knowledge, the Applied Ontology course focuses on how to structure the content of a knowledge base and introduces general structuring principles for knowledge. Examples include the theories for mereological (parthood) relations, or theories of space and time, and the consequences of selecting a particular theory in formalized knowledge bases. The course is split in two parts, the first focusing on concrete applications with examples taken from the biomedical domain, the second focusing on the theoretical framework underlying formal ontologies and their role in information systems.
Course Description from Program Guide	The course covers advanced topics in conceptual modelling, data management, integration and analysis, all of which have applications in data-intensive disciplines such as biology, biomedicine and others. The aims of the course is to provide an in-depth understanding of the state of the art in formal ontologies, including their role in integrating and analyzing data. While Knowledge Representation and Reasoning (CS213) introduced basic logic formalisms that can be used to express knowledge. Examples include the theories for mereological (parthood) relations, or theories of space and time and the consequences of selecting a particular theory in formalized knowledge bases. The Course is split in two parts, the first focusing on concrete applications with examples taken from the biomedical domain, the second focusing on the theoretical framework underlying formal ontologies and their role in information systems.
Goals and Objectives	The aims of the course is to provide an in-depth understanding of the state of the art in formal ontologies, including their role in integrating and analyzing data. At the end of the course, students will be able to apply methods of ontological analysis and modelling to generate formal knowledge bases and use ontology-base approaches such as semantic similarity measures to analyze data.
Required Knowledge	Knowledge representation and Reasoning (CS213), Data Analytics

Reference Texts	Knowledge Representation: Logical, Philosophical, and Computational Foundations. John Sowa, Brooks Cole Publishing Co, 1995. Formal Ontology and Information Systems. Nicola Guarino, Proceedings of Formal Ontology and Information Systems, IOS Press, 1998. A Catalog of Temporal Theories. Pat Hayes, Beckman Institute and Departments of Philosophy and Computer Science, 1996. General Formal Ontology (GFO) - A Foundational Ontology Integrating Objects and Processes. Herre, Heller, Burek, Hoehndorf, Loebe, Michalek, Onto-Med Report, 2006. WonderWeb Deliverable 18: Ontology Library. Masolo, Borgo, Gangemi, Guarino, Oltramari, 2003. An overview of OntoClean. Nicola Guarino, Christopher Welty. Handbook on Ontologies, Springer, New York, 2004. The role of ontologies in biological and biomedical research: a functional perspective. Hoehndorf, Schofield, Gkoutos, Briefings in Bioinformatics, 2015. Semantic Similarity in Biomedical Ontologies. Pesquita, Faria, Falco, Lord, Couto, Plos Computational Biology, 2009.
Method of evaluation	**30.00**% - Presentation **70.00**% - Research Project
Nature of the assignments	The final grades will be based on one research-type project in which an ontology is designed, data from a domain of knowledge is integrated and analyzed using this ontology. The project will be presented at the end of the course, and a written report should be prepared describing the project. The report may form the basis for a research paper. Additionally, two presentations of research papers will contribute to the final grades.
Course Policies	The course project must be finalized and the final report submitted by approximately 1 week before the end of the semester. In exceptional cases, additional time may be permitted.
Additional Information	

Tentative Course Schedule *(Time, topic/emphasis & resources)*		
Week	**Lectures**	**Topic**
1	Sun 08/26/2018 Wed 08/29/2018	Introduction to biomedical ontologies: We will introduce ontologies as they are applied in biology and biomedicine. We will focus in particular on the medical ontology GALEN and the Gene Ontology.
2	Sun 09/02/2018 Wed 09/05/2018	Representation formats for biomedical ontologies: We will discuss formal languages to represent biomedical ontologies, including the OBO Flatfile Format and the Web Ontology Language (OWL), as well as their interrelations. We will also discuss approaches for generating graph representations of ontologies.
3	Sun 09/09/2018 Wed 09/12/2018	Ontology-based data integration: We will discuss how ontologies are used for integration of data in biology, and the role of deductive inference in integrating datasets. We will investigate the role of OWL reasoning and different OWL profiles that enable tractable reasoning so that large ontologies can be integrated efficiently.
4	Sun 09/16/2018 Wed 09/19/2018	Statistical analysis of data with ontologies: We will introduce the Gene Set Enrichment Analysis method which is widely used for the interpretation of gene expression datasets. We will demonstrate its origins in the Gene Ontology and how to apply it in ontologies of other domains.
5	Sun 09/23/2018 Wed 09/26/2018	Semantic similarity: In this lecture, we introduce and discuss semantic similarity measures defined over ontologies, and their role in the analysis of data. We show that ontology-based similarity can identify functional similarity between gene products and that this similarity can be used to identify interactions between proteins. We will then generalize the result and discuss the conditions under which semantic similarity can be applied to identify relations between entities characterized through ontologies, including the application of semantic similarity in natural language processing.
6	Sun 09/30/2018 Wed 10/03/2018	Data-driven generation of ontologies: Generation of ontologies is often a manual process that does not easily scale with the increase in data and knowledge we are faced with in biology. We will discuss methods to automatically generate ontologies, starting with approaches relying on natural language processing, and specifically focus on approaches that use empirical data and measure- ments to generate ontologies of phenotypes and cellular locations from high-throughput datasets. We generalize these results and introduce ontology design patterns.
7	Sun 10/07/2018 Wed 10/10/2018	Introduction to formal ontologies in information systems: This lecture will give a bird's-eye perspective on the field of formal ontology. We will discuss history of ontology, it's role in computer science, and current research directions. Reading material: Guarino (1998): Formal Ontology and Information Systems.
8	Sun 10/14/2018 Wed 10/17/2018	Mereology: We will introduce different theories of parthood, including meta-theoretical properties such as completeness and decidability. We will discuss parthood as a partial order, and extensions such as strong and weak supplementation principles, existence of largest or smallest elements, atomicity, as well as extensionality. We will apply different theories of mereology to define collections and collectives.

9	Sun 10/21/2018 Wed 10/24/2018	Ontology of space and time: We will introduce different theories to represent time, based on Hayes' catalog of temporal theories, as well as different theories for representing space. We will discuss key limitations of the naive representation of time using real numbers and develop alternatives.
10	Sun 10/28/2018 Wed 10/31/2018	Ontology of space and time (cont.)
11	Sun 11/04/2018 Wed 11/07/2018	Top-level ontology: For two weeks, we will introduce top-level ontologies, including the Descriptive Ontology for Linguistic and Cog- nitive Engineering (DOLCE), the General Formal Ontology (GFO), and the Basic Formal Ontology (BFO). We will distinguish endurantism and perdurantism and investigate how these are represented in the different top-level ontologies, as well as the theories of time underlying them.
12	Sun 11/11/2018 Wed 11/14/2018	Top-level ontology (cont.)
13	Sun 11/18/2018 Wed 11/21/2018	Ontology of properties and OntoClean: We will discuss in depth the OntoClean methodology and its associated ontology of properties as a method to formally verify and evaluate a formal ontology. We will use examples from top-level ontologies as well as biological domain ontologies to demonstrate the utility of the method.
14	Sun 11/25/2018 Wed 11/28/2018	Functions, roles and relations: The lecture will deepen the discussion of key components of ontologies, focusing on functions, roles and relations. A particular focus will be on qua-individuals and the notion of intentionality.
15	Sun 12/02/2018 Wed 12/05/2018	Project presentations
16	Sun 12/09/2018	

Note

The instructor reserves the right to make changes to this syllabus as necessary.