Semantic Web and Ontologies: Lab 2

OWL Ontology Building

This lab is to help to get you familiar with the process of defining an ontology for a domain. Ontology is an important concept for Semantic Web, for which W3C has published OWL and OWL 2 as the standard languages. An ontology is to encode pieces of background knowledge for a domain, and they are saved in the format .owl, guaranteeing a great information sharing among different applications. The description logics based semantics equips OWL and OWL2 with a sound unambiguous meaning to prevent mis-interpretations among different users of these data. Note that, we use the terminologies "property", "role", and "attribute" interchangeable to mean a binary relation.

You are asked first to write down several elements of an ontology in paper. You are then assigned to build an ontology with the tool Protégé (or OWL API) and check the inconsistency of your ontology with the plugin reasoner of Protégé (or in OWL API).

Exercise 1. Browse an ontology (with Protégé and/or OWL API)

1. Download Protégé from `http://protege.stanford.edu/`.
2. Run Protégé from the terminal: path-protege/run.sh
3. Import an existing ontology into Protégé (e.g. the pizza ontology `https://protege.stanford.edu/ontologies/pizza/pizza.owl`
4. Verify the ontological elements (concepts, properties, instances, axioms) in the ontology
5. Compare the tabs "Class hierarchy" and "Class hierarchy (Inferred"). Are they different? Why?

Exercise 2. Reasoning with your ontology by a reasoner With Protégé, choose a reasoner by clicking on the button "reasoner", and then "start reasoning".

- Is your ontology consistent?
- What do the buttons "Stated" and "Inferred" mean?

Exercise 3. This exercise is to recall that description logics is not a single logic, but a family of languages, among which is $\mathcal{ALC}$ as we discussed in the lecture. Indeed, OWL and OWL 2 standards contain 7 different description logic languages that differ in syntax and expressivity.

Consider the university study scenario, in particular an Examination Regulation given in the Appendix, and answer the following questions.

- List the atomic concept names, role names, and individual names for this scenario. (For example, $Students$ is an atomic name, $hasStudentId$ is a role, and $Pierre$ can be an individual)

- Construct several complex concepts using different Description Logic constructors given below:
$$\sqcup, \sqcap, \neg, \exists\ R.C, \forall\ R.C, \geq n\ R.C, \leq n\ R.C$$
For each of these complex concepts, give its explanation in natural language.
For example, $Student \sqcap \exists hasDiploma.Engineering$ means "all students who have obtained Engineering diploma".

- Construct a TBox and a ABox for this domain. In Description Logic, an element from TBox or Abox is called an axiom.
For example, $Apprentice = Student \sqcap \exists\ WorkIn.Enterprise$ can be an element from TBox (so it is called an axiom), and $WorksIn(John, Thales)$ can be an ABox element.
You are required to build a TBox having at least 3 TBox axioms and an ABox having at least 3 axioms.

- Can you infer some implicit information from the ontology $\mathcal{O}$ =(TBox, ABox) from the TBox and ABox you built above?
For example, I can deduce $Apprentice(John)$ based on the TBox
$$\mathcal{T} = \{Apprentice = Student \sqcap \exists\ WorkIn.Enterprise\}$$
and ABox
$$\mathcal{A} = \{WorksIn(John, Thales), Student(John), Enterprise(Thales)\}.$$

Exercise 4. Creation of an ontology (with Protégé and/or OWL API)

1. Create a new OWL ontology with Protégé that encodes the answers you have for Excise 3. And try to enrich it (as much as possible) to model the knowledge conveyed by the exam regulation.

2. The creation of OWL ontology is also doable via OWL API.

Exercise 5. (Optional) This exercise is to understand a typical ontology browser for the biomedical ontology Snomed CT.
Go to the webpage `http://browser.ihtsdotools.org`, and click "Go browsing International edition" to enter the latest version of Snomed CT browser page. On the left, you can see the list of concepts defined in this ontology, by clicking on a concept, you will get information about it in the right column. You can search a concept by its id, and you can click on "Summary" and "Diagram" to figure out its information. Can you write down the description logic formats for the following concepts?

- 131148009(Bleeding),

- 844005(Behavior finding),
- 420395004

For example, $262832002 = 30171000 \sqcap \exists\ 363698007.23451007 \sqcap \exists\ 116676008.35566002$.

Note that 284666000(Trophic life form) was once in the ontology, but not anymore the case. So ontology is evolving.

Appendix

Students The admission jury decides which of the students who apply for the program are admissible. One condition for being admissible is having completed at least 4 years of ternary study in Computer Science or a related field (in France: Bac+4). An admissible candidate becomes a student of the program, if he/she has (1) registered in the proper Paris-Saclay establishment (see Web Page of Paris-Saclay University), (2) registered with the study inspector, and (3) paid any associated fees. All of this has to happen for the current year of study, and before the second week of the start of the current year of study. Exceptions can be granted by the study coordinators in agreement with the reference institute and study coordinator. The study inspector maintains the list of students. If a student does not participate regularly in the program for one period, their inscription can be revoked. The rules that the reference institute has put in place for its diploma program apply accordingly.

Offered Courses A course is a unit of instruction (typically a sequence of teaching hours, labs, and/or projects) organized by one or several lecturers at Paris-Saclay University. The program encompasses

- A number of courses called "mandatory"
- A number of courses called "optional"

The courses and their ECTS credit numbers are proposed by the program coordinators, in agreement with the lecturers. The courses are displayed on the Web page of the program. All obligatory courses will necessarily take place. The optional courses are not guaranteed to take place (they may or may not start, depending on the number of interested students, and other factors). One optional course is the "Module Liberté". It is a placeholder for any data-oriented course at Paris-Saclay University, subject to the approval of the program coordinators. A student can follow at most one Module Liberté. The Module Liberté will necessarily take place. The Module Liberté is worth a predetermined number of ECTS credits, no matter how many credits are awarded by the program that offers the course. Courses typically take place in the first semester of the study year. Students that are not registered for the program can participate in the courses at the discretion of the lecturer.

Examinations A course is evaluated by examinations. An examination can be a midterm exam, a final exam, projects, homework, graded practical labs, presentations or similar items. Each examination is graded on a 0-20 scale, with 20 being the maximal grade and 10 being the passing grade. The following applies to all courses except for the Module Liberté. Each

course offers two rounds of examinations. Exams of both rounds take place during predetermined periods. For each round, the lecturer defines a grading scheme, i.e., (1) the examinations that play into the grade of the round and (2) the weight of these examinations for the grade of the round. The grading scheme of the second round can include examinations from the first round. The lecturer communicates the grading schemes to the program coordinators before the start of the study year. The grading schemes are displayed on the Web page of the program at most one month after the start of the study year. If a student commits fraud in an examination (including plagiarism), the grade of this examination is zero. If the student copied from another student in the course, the other student also receives a grade of zero for this examination. Upon request by either the students or the lecturer, the case is referred to the jury. The jury can either exonerate the students or ask the reference institute to apply the sanctions for fraud that it has put in place for its diploma program.