CS 227: Homework #2

Assigned: Tuesday, April 19th
Due:
<u>Tuesday, April 26th:</u> Problem 2(a) and 2(b) to be completed online
<u>Thursday April 28th:</u> All four problems to be submitted in class or by 5pm in the CS227 file cabinet in the first floor lobby.

Please note this assignment requires more work than your previous assignment. We recommend that you start early so you can give yourself enough time to finish. As always, please email us if you have any questions or you are unsure about a specific part.

Problem 1: Applying Structured Descriptions

Goal

The goal of this problem is to expose you to the task of modeling real world knowledge by applying the knowledge representation concepts of object oriented representation, description logics, subsumption reasoning and inheritance. We will do this by representing one paragraph of knowledge about the human heart. This problem is a miniature version of a larger problem of representing textbooks and answering questions at the back of the book.

Overview of the problem

The material that we will use is freely available on Wikipedia at http://en.wikipedia.org/wiki/Human_heart and we include it here for easy reference:

The **human heart** is a muscular organ that provides a continuous blood circulation through the cardiac cycle and is one of the most vital organs in the human body.[1]The heart is divided into four main chambers: the two upper chambers are called the left and right atria and two lower chambers are called the right and left ventricles.There is a thick wall of muscle separating the right side and the left side of the heart called the septum. Normally with each beat the right ventricle pumps the same amount of blood into the lungs that the left ventricle pumps out into the body. Physicians commonly refer to the right atrium and right ventricle together as the **right heart** and to the left atrium and ventricle as the **left heart**.[2]

You will do this in three steps:

- Design the knowledge representation on paper.
- Encode the knowledge that you generated into OWL using a knowledge base environment called Protégé
- Use the built in reasoner Pellet to demonstrate three key description logic inferences
  - Check that the KB you created is consistent
  - Classify the taxonomy
  - Create a sample individual and infer its types

To get you started, a complete step-by-step worked out example process for the first sentence is included in this handout. Your task is to represent the remaining sentences by defining any necessary concepts, relations and individuals, and writing their definitions as done for the first sentence in the worked out solution. We have also provided the sample OWL file for the first sentence that you can load into Protégé.

Protégé Setup

1. Go to the Protégé website at http://protege.stanford.edu/
2. From the top menu choose Download and as a registered user continue to the download page
3. We will be using version 3.4.5 of Protégé available from: http://protege.stanford.edu/download/protege/3.4/installanywhere/
   Please download and install the tool. A tutorial on how to use it is available at:

http://owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/resources/ProtegeOWLTutorialP3_v1_0.pdf

Knowledge Representation and Reasoning Steps

1. Write out the structured descriptions for the concepts that appear in the paragraph.
   a. This includes defining classes, relations, individuals, and writing their definitions as done in the sample provided here.
   b. You can use the description language syntax as defined in the Chapter 9 of the B&L textbook.
   c. You can use any concept or role forming constructors that are available in the OWL specification that you have already studied as part of HW1.

2. Encode the structured descriptions using the Protégé ontology editing tool in the OWL plugin.
3. Under the reasoning menu, select Pellet 1.5.2 as the reasoner. Once selected, perform the following operations:
    a. Check consistency
        i. If during consistency checking, some errors are found, you would need to fix those errors before you can perform the remaining steps.
    b. Classify taxonomy
    c. Compute inferred types

Submission

You need to submit:

1. Your paper design of the knowledge base for the paragraph that specifies the structured descriptions for concepts that appear in the paragraph
2. Paper documentation of the knowledge base that you encoded that includes at least the following information
    a. Screen shot of the class taxonomy that you created
    b. A screen shot of the taxonomy after you classify the taxonomy
    c. A screen shot showing the inferred types of the new individual that you created.
3. A summary giving new super classes and types that were computed and any problems or challenges that you faced during the process.

As an illustration of the process, a complete worked out solution for the first sentence follows. The methodology used here is very close to what was described in the article entitled *Ontology Development 101* that was an assigned reading for the lecture on *Ontologies*.

Example Solution for the First Sentence

1. *Scope of the knowledge to be modeled:*

For this exercise, we will represent the knowledge necessary to represent the first sentence:

The **human heart** is a muscular organ that provides a continuous blood circulation through the cardiac cycle and is one of the most vital organs in the human body.[1]

Enumerate important terms:

The above sentence involves the following terms

- human
- heart
- muscular
- organ
- blood
- circulation
- cardiac cycle
- vital
- body
- provides
- through

Define the classes and the class hierarchy

- We define the most general class as *Thing,* and define its two subclasses *Entity* and *Event.*
- We define the class *Human, Heart, Blood, Organ, Muscle,* and *Body* as a subclass of *Entity*, and define *Circulation* and *Cardiac Cycle* as the subclasses of *Event.* We define a general class *Physical-Object* as a subclass of *Thing*.
- We define a class *Quantity* as a subclass of *Thing*. We define its subclass *Importance-Value.*

This gives us a class taxonomy shown below:

```
Thing
        Entity
                Human
                Heart
                Blood
                Body
                Organ
                Muscle
        Event
                Circulation
                Cardiac Cycle
        Quantity
                Importance-Value
```

Define the relations

- We define the relation *is-part-of* with a domain of *Entity* and a range of *Entity.* (You have already learned about domain and range constraints as part of Homework 1.) We define *has-part* as an inverse of the relation *is-part-of.*

- We define a relation *material* with a domain of *Entity* and range of *Entity*.
- We define a relation *is-possessed-by* with a domain of *Entity* and range of *Entity*. The inverse of *is-possessed-by* is *possesses*.
- We define a relation *importance* with a domain of *Entity* and a range of *Importance-Value.*
- We define a relation *object* with a domain of *Event* and a range of *Entity*.
- We generalize the notion of *provides* and define a relation *has-function* with the domain *Entity* and *range* of *Event.*
- We generalize the notion of *through* and define a relation *by-means-of* with a domain of *Event* and a range of *Event.*

Define the Individuals

- We define three individual instances of Importance-Value: *Low, Medium, High.*

Structured Descriptions for the question and the answer

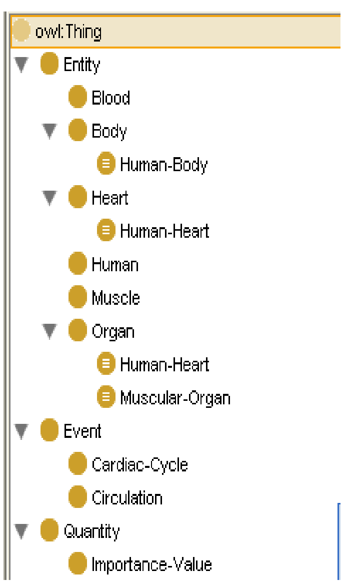
The structured descriptions in this section are written using the notation that was introduced in the Chapter 9 of the B&L textbook with one exception: we have used the symbol ≡ to assign a name to a complex concept.

(Entity ⊆ Thing)
(Event ⊆ Thing)
(Quantity ⊆ Thing)
(Human ⊆ Entity)
(Heart ⊆ Entity)
(Blood ⊆ Entity)
(Body ⊆ Entity)
(Organ ⊆ Entity)
(Circulation ⊆ Event)
(Cardiac-Cycle ⊆ Event)
(Importance-Value ⊆ Quantity)
Low → Importance-Value
Medium → Importance-Value

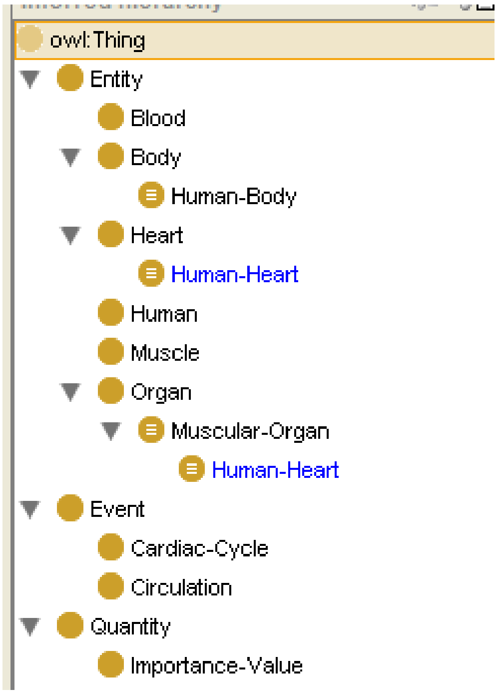
High → Importance-Value

(Muscle ⊆ Entity)
(Human-Heart ≡ [**AND** Heart
[**SOME** is-part-of Human]])
(Human-Heart ⊆ [**AND** Heart Organ
[**SOME** material Muscle]
[**SOME** has-function
[**AND** Circulation
[**SOME** object Blood]
[**SOME** by-means-of Cardiac-Cycle]]])
(Muscular-Organ ≡ [**AND** Organ
[**SOME** material Muscle]])
(Human-Body ≡ [**AND** Body
[**SOME** is-possessed-by Human]])
(Human-Body ⊆ [**AND** Body
[**SOME** has-part [AND Human-Heart
[FILLS importance High]]]])
Heart-30→Heart
John → (Human [FILLS has-part Heart-30])

The following screen shot shows the class taxonomy before computing the classification.

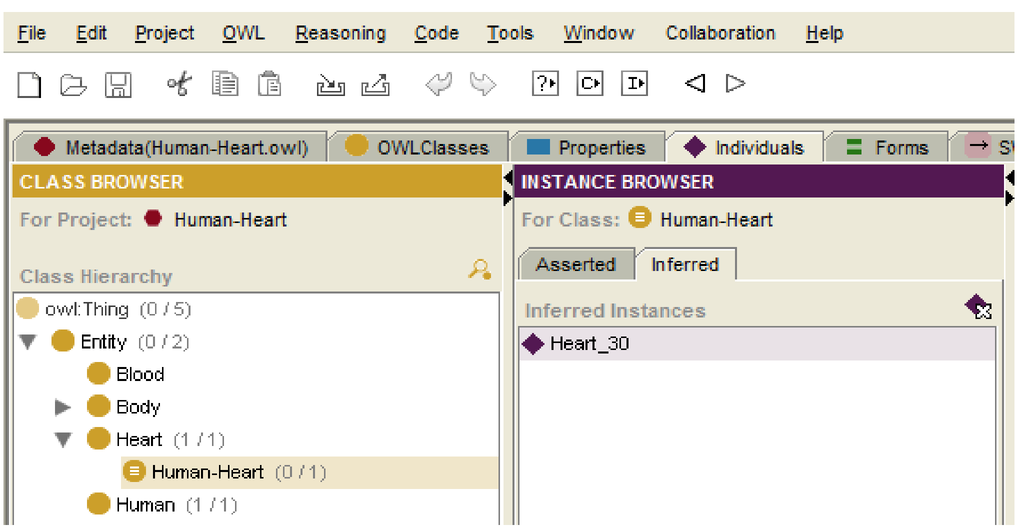


The following screen shot shows the class taxonomy after computing the classification.



Human-Heart was automatically classified under Muscular-Organ.

The following screen shot shows that Heart-30 was automatically inferred to have a type of Human-Heart.



Problem 2: Knowledge Representation in Social Context

In this problem, we will experiment with a Semantic Media Wiki (SMW) as a platform for doing knowledge representation and reasoning in a social context. The goal is to understand how doing KR&R in open social context is different than what we did in Problem 1.

We will use the SMW available at http://wiki.projecthalo.com. This installation has been made generously available to us for the purpose of this course by Vulcan Inc. You can go to this site and create a user account for yourself for the purposes of this assignment.

The social context that we will use for this assignment will be defined by the students of this class. Working together, the class will create a Wiki about Cars. To get started, it will be helpful for you to read the Tutorial to Media Wiki which is linked from the Wiki home page. The Wiki has been pre-populated with a simple ontology about Cars. You can view the existing ontology by using the Ontology Browser. The Wiki also has pre-built semantic forms to create a new car company, a new car model, and a new individual instance of a car which you will see on the right hand side of the Wiki home page.

a) Using the semantic form available on the right hand side of the Wiki home page, create an individual wiki page that represents your car. This can be your personal car, or your favorite car, or your dream car, or your friend's car, etc. Before creating this individual instance of a car, make sure that you have defined an appropriate model and manufacturer for this car. Since the wiki is pre-populated with only Audi and BMW as manufacturers, and a few models, it is unlikely that the manufacturer and the model for the individual car you want to create already exists. The manufacturer and the model could have been also created by one of your class mates, and if that is the case, re-use it in favor of creating one from scratch. As part of the description of your car, use the properties *has-interior-color* and *has-exterior-color* to define its interior and exterior colors. Here is an example description of the car:

Vinay's car is [[has-exterior-color::Black]] and has an interior color of [[has-interior-color::Brown]].

Once you have created the markup, you should be able to see the RDF corresponding to it by clicking on the ``View RDF'' link on the page.

You can see several additional sample Wiki entries by clicking on the pages under ``Audi Pages'' on the home page of the wiki.

b) Expand the description of your car as a wiki page entry. The page should be at least one paragraph, but there is no upper limit. You can write whatever you find interesting, add photographs, link to other cars, interesting embedded queries, etc. Your wiki page should include at least three properties that you believe will be generally useful to the class. These properties could be either created by you or by others in the class. For example, you might include properties of cars like *has-seating-capacity* or *has-body-style*. Part of the power of doing KR in a social context is for a group to converge on a set of appropriate classes and properties. For example, your expanded wiki entry could look like:

This is a two door car with [[has-seating-capacity::four]] seats but is ideally suited for only two people.

Vinay bought this car back in 2006 inspired by Mark Greaves. [[liked-by::Mark Greaves]] loves convertibles and wanted to enjoy this vicariously through Vinay's purchase. One day Mark came to SRI for a business meeting, but after he learned about the new purchase, he went for a drive to [[favorite-driving-destination::Skyline Boulevard]]. Even though it was a bit foggy that day, it was very clear at the ridge top, and was a perfect day to go driving top down.

This car [[Model::325ci]] is owned by [[Owner::User:chaudhri]]

The last sentence above is necessary to update the *Total Ownership* pie chart on the home page of the Wiki.

c) On Tuesday, April 26th, at 5PM, the wiki will be frozen for writing. Once it is frozen, analyze the socially created markup of the wiki page entries. Specifically, address the following questions. Your report addressing the final points will be the only formal submission required for this part.

- A link to the Ask interface for the Wiki is available from the home page. Use the Ask interface to create statistics pages which contain at least the following queries:
    1. i. Which people in the class have a car with the same manufacturer? List the people and the manufacturer.
    2. ii. Which people in the class have a car with the same exterior color? List the people, car, and the color.
    3. iii. Which is the most popular car in the class?
- Did common properties emerge that were used by several people?
- Did people define properties that were identical in meaning but differently named?
- What interesting queries about the wiki entries are possible because they are marked up using formal representation?

d) The final part is for extra credit and our incentive to the class. The extra credit will be given depending on how interesting the final outcome of the socially created markup of the wiki entries is. You can also create a social process with some mutual incentive scheme so that the pages include common themes: for example, favorite driving destinations. Either the whole class will get the extra credit or no one will get the extra credit. The judgment for the extra credit will be made by the course staff. This is an opportunity for the class to create social processes and incentive schemes to create an artifact that is of most interest to you as well as to others.

Problem 3: Logic Programming

(This problem corresponds to the first three parts of the first problem at the end of Chapter 6 in the B&L textbook)

Assume that a KB consists of a list of rules of the form $(q \leftarrow a_1 \ldots\ldots a_n)$ where $n \geq 0$, q is an atom, and each a_i is either of the form p or **not**(p), where p is an atom. The q in this case is called the conclusion of the rule, and the a_i's make up the antecedent of the rule.

The forward-chaining procedure presented in Chapter 5 for Horn clause satisfiability can be extended to handle negation as failure by marking atoms incrementally with either a Y (when they are known to be solved), or with an N (when they are known to be unsolvable), using the following procedure:

For any unmarked atom q,

- if there is a rule $(q \leftarrow a_1 \ldots\ldots a_n) \in$ KB, where all the positive a_i are marked Y and all the negative a_i are marked N, then mark q with Y.
- if for every rule $(q \leftarrow a_1 \ldots\ldots a_n) \in$ KB, some positive a_i is marked N or some negative a_i is marked Y then mark q with N.

Note that the first case trivially applies for rules where $n = 0$, and that the second case trivially applies if there are no rules with q as the conclusion.

(a) Show how the procedure would label the atoms in the following KB:

$a \leftarrow$
$b \leftarrow a$
$c \leftarrow b$
$d \leftarrow$ **not**(c)
$e \leftarrow c, g$
$f \leftarrow d, e$
$f \leftarrow$ **not** (b) $, g$
$g \leftarrow$ **not**(h), **not**(f)

(b) Give an example of a KB where this procedure fails to label an atom as either Y or N, but where the atom is intuitively Y according to negation as failure.

(c) A KB is defined to be strongly stratified if and only if there is a function f from atoms to numbers such that for every rule ($q \leftarrow a_1 . . . . . a_n$) $\in$ KB, and for every $1 \leq i \leq n$, it is the case that $f(q) > f(a_i)$, where $f(\textbf{not}(p_i)) = 1+f(p_i)$. (In other words, the conclusion of a rule is always assigned a higher number than any atom used positively or negatively in the antecedent of the rule.) Is the example KB of part (a) strongly stratified?

Problem 4: Logic Programming

(This problem is the same as the first three parts of the problem 4 at the end of Chapter 5.)

In this question, we will explore the semantic properties of propositional Horn clauses. For any set of clauses *S*, define $\mathtt{I}_S$ to be the interpretation that satisfies an atom *p* if and only if $S \models p$.

(a) Show that if *S* is a set of positive Horn clauses, then $\mathtt{I}_S \models S$.
(b) Give an example of a set of clauses S where $\mathtt{I}_S \not\models S$.
(c) Suppose that S is a set of positive Horn clauses and that *c* is a negative Horn clause. Show that if $\mathtt{I}_S \not\models c$ then $S \cup \{c\}$ is unsatisfiable.