UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

B. Sc. Examination 2010

Computing and Information Systems

IS53020A (CIS344)   The Semantic Web

Duration: 2 hours 15 minutes

Date and time:

There are five questions in this paper. You should answer no more than three questions. Full marks will be awarded for complete answers to a total of three questions. Each question carries 25 marks. The marks for each part of a question are indicated at the end of the part in [.] brackets.

There are 75 marks available on this paper.

No calculators should be used.

**THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM**

Question 1   Triplestores and graphs

Restaurant	Address	Cuisine	Licensed	Price
Peking Garden	Preston St	Chinese	Y	££
Grubb	St James St	Burgers	N	£
Regency	Marine Drive	Fish	Y	£££
King of Siam	High St	Thai	Y	££

Guide to local restaurants

(a) Explain what is meant by a **triplestore** in the context of the semantic web, contrasting it with relational databases and spreadsheets.   [3]

(b) The table above, **Guide to local restaurants**, lists various facts about restaurants in tabular form. How would you express these facts in the form of a triplestore? Explain your answer.   [9]

(c) What do you consider would be some advantages and disadvantages of a triplestore over the tabular format? Support your argument with appropriate examples.   [4]

(d) Suppose you were asked to extend the triplestore to include facts about cinemas in the locality, as in this table:

Cinema	Address	Screens	Company	Price
Brighton Odeon	West St	8	Odeon	££
Duke of Yorks	London Road	1	Picturehouse	£
Cinematheque	Middle St	1	Independent	£

What extra lines would you need to add? Justify your answer.   [4]

(e) Show how your new triplestore can be represented as a graph or a series of graphs. Explain how you have constructed the graph(s).   [5]

Question 2   Ontologies

In this question, you should assume that names of properties and classes such as **Houses**, **teaches** etc have the usual meanings of these words in ordinary English.

(a) i. Briefly explain what is meant by an **ontology** in computer science and give some reasons why ontologies have an important role to play in the development of the Semantic Web.

   ii. Explain the difference between **top-down** and **bottom-up** modelling in ontology development.

[5]

(b) i. Suppose an ontology includes the classes listed below. Which of them are **sub-classes** or **super-classes** of the **Buildings** class? List all that apply and justify your answers.

   Rooms
   Windows
   Churches
   Structures
   Corridors
   Houses
   Sheds
   Mosques
   Bricks
   Architects

   ii. Suppose the ontology includes a property **teaches**, as in *Professor Winter* **teaches** *Mathematics*. Which of the following statements involve **sub-properties** or **super-properties** of the **teaches** property? List all that apply and justify your answers.

   Professor Winter **teaches** calculus.
   Professor Winter **sets exams in** Mathematics.
   Professor Winter **lectures in** Mathematics.
   Professor Winter's research assistant **teaches** Mathematics.
   Professor Winter **gives tutorials in** Mathematics.
   Professor Winter **is learning** Mathematics.
   Professor Winter **imparts knowledge of** Mathematics.
   Professor Winter **is an expert** in Mathematics.

[10]

(c) "An ontology together with a set of individual instances constitutes a knowledge base. In reality, there is a fine line where the ontology ends and the knowledge base begins". (Noy and McGuiness, Ontology Development 101)

What considerations could determine whether the following are specified as classes, properties or instances in a knowledge base for a wine ontology?

  i. Wine regions: e.g. North/South American, French, Beaujolais, Napa Valley etc.

 ii. Wine varieties: Jacobs Creek Cabernet Sauvignon, Sterling Vineyards Merlot etc

iii. Wine colours: red, white, rose

[10]

Question 3   Semantic Web Languages

(a) i. Antoniou and van Harmelen (2004) state that "XML allows the representation of information that is also machine-accessible" (p. 25) but also that XML in comparison with RDF "does not provide any means of talking about the semantics (meaning) of data" (p. 61). How can these statements be reconciled?

ii. Explain in general terms the relationship between RDF and RDF Schema.

[7]

(b) The following is a slightly simplified extract from an RDF document in triples format.

```
<http://mydom/uni#prof_5> <http://mydom/uni#name> "Jim Dixon" .


<http://mydom/uni#prof_5>
 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
 <http://mydom/uni#professor>   .


<http://mydom/uni#academic_staff>
 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
 <http://www.w3.org/2000/01/rdf-schema#Class> .


<http://mydom/uni#professor>
 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
 <http://www.w3.org/2000/01/rdf-schema#Class> .
```

i. Express the content of the document in ordinary English.

ii. Suppose you wanted to implement a range restriction to the effect that courses must be taught by academic staff only. Would Jim Dixon be allowed to teach courses? If not, what would have to be added to the document for him to be allowed to? Justify your answer.

[8]

(c) i. Explain what is meant by transitive, symmetric and functional properties in the OWL language.

ii. The following is an excerpt from an OWL document, coded in the Manchester OWL Syntax. Explain the use of the **subClassOf** property, and express the content in ordinary English:

```
Class: AcademicStaffMember
    SubClassOf:
        StaffMember,
        teaches some UnderGraduateCourse
```

[10]

Question 4   RDF and SPARQL

The following is an excerpt from an RDF document in the Turtle format. The document includes **class** and **property** definitions but no individual instances. Classes are coded as RDF Schema elements and properties as OWL Object or Datatype Property elements.

```
@prefix xsd:      <http://www.w3.org/2001/XMLSchema#> .
@prefix default:  <http://www.doc.gold.ac.uk/~mas01rk/Movies/#> .
@prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#> .
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix owl:      <http://www.w3.org/2002/07/owl#> .

default:title
      rdf:type owl:DatatypeProperty ;
      rdfs:domain default:Films ;
      rdfs:range xsd:string .
<http://www.doc.gold.ac.uk/~mas01rk/Movies/>
      rdf:type owl:Ontology .
default:name
      rdf:type owl:DatatypeProperty ;
      rdfs:domain default:People ;
      rdfs:range xsd:string .
default:appearsIn
      rdf:type owl:ObjectProperty .
default:Films
      rdf:type rdfs:Class .
default:starsIn
      rdf:type owl:ObjectProperty ;
      rdfs:subPropertyOf default:appearsIn .
default:genre
      rdf:type owl:DatatypeProperty ;
      rdfs:domain default:Films ;
      rdfs:range xsd:string .
default:year
      rdf:type owl:DatatypeProperty ;
      rdfs:range xsd:int .
default:People
      rdf:type rdfs:Class .
default:directed
      rdf:type owl:ObjectProperty ;
      rdfs:domain default:People ;
      rdfs:range default:Films .
```

(a) Draw a graphical representation of the following set of facts and instances, based on the class and property definitions specified in the above document: [8]

Instances of the class Films are Psycho (1960), The Trial (1962), Catch-22(1970) and Rocky Horror Show (1975).

Instances of the class People are Anthony Perkins and Alfred Hitchcock.

Anthony Perkins stars in Psycho and The Trial and appears in Catch-22.

Alfred Hitchcock directed Psycho.

Psycho belongs to the genre Horror.

Rocky Horror Show belongs to the genres Horror, Comedy and Musical.

(b) Using the SPARQL Query Reference included with this exam paper, construct queries to retrieve the following information (assuming the individual instances have been appropriately defined): [12]

  i. What year was Psycho released?
  ii. List all films directed by Alfred Hitchcock and starring Anthony Perkins.
  iii. List the titles and dates of all films directed by Steven Spielberg in alphabetical order.
  iv. List Hitchcock's five earliest films.

(c) What are the three main components of the SPARQL specification? What are some limitations of the SPARQL language? [5]

Question 5   RDF and SPARQL: Applications

(a) Write brief notes on **one** of the following organisations' use of Semantic Web technologies:

    i. Google

   ii. Oracle

  iii. The BBC

[5]

(b) i. Describe the main characteristics of the FOAF RDF vocabulary. What purposes was the vocabulary originally designed for, and how is it frequently used outside the original domain?

  ii. Describe another example of an RDF vocabulary which has been taken up by the wider modelling community.

[8]

(c) i. Explain what is meant by a **SPARQL endpoint**, and the difference between **generic** and **specific** endpoints.

  ii. Give a detailed description of a specific SPARQL endpoint resource you have encountered during this course (up to about 500 words).

[12]