Two hours – online

EXAM PAPER MUST NOT BE REMOVED FROM THE EXAM ROOM

NANJING UNIVERSITY

SCHOOL OF ARTIFICIAL INTELLIGENCE

KNOWLEDGE REPRESENTATION AND PROCESSING

Date: Monday, 27 July 2020

Time: 14:00 - 16:00

This is an offline examination.
The examination contains SHORT ESSAY QUESTIONS.
Be sure to answer ALL Questions.

This is an OPEN book examination

The use of electronic calculators is NOT permitted

1. This question is on knowledge representation and AI.

✧ In 1993, Gruber originally defined the notion of an ontology as an "explicit specification of a conceptualization". In 1997, Borst defined an ontology as a "formal specification of a shared conceptualization". This definition additionally required that the conceptualization should express a shared view between several parties, a consensus rather than an individual view. Also, such conceptualization should be expressed in a (formal) machine readable format. In 1998, Studer et al. merged these two definitions stating that: "An ontology is a formal, explicit specification of a shared conceptualization."

Do you think this is a good definition for ontology? Why and why not?

(3 marks)

2. This question is on knowledge acquisition and modelling.

✧ Techniques to perform ontology development range from manual to almost fully automated. In what way(s) may data mining be useful in ontology development? Your answer should include something about the following three aspects:

a. populating the TBox (learning classes and hierarchies, relationships, constraints),

b. populating the ABox (assertions about instances), and

c. possible substitutes or additions to the standard automated reasoning service (consistency checking, instance classification, etc.)

(6 marks)

✧ Consider again the scenario from the previous question:
Animals

- fall in different categories such as Mammals, Birds, Reptiles, Insects, Fish, etc.
- live in different habitats such as forest, desert, ocean, river, etc. These habitats can be in different geographical regions, including countries or continents.
- eat something, for example plants or other animals.
- move (with very few exceptions), for example by walking or swimming or flying.

They usually do this by making use of certain body parts, for example wings, feet, or hoves. Assume you want to describe, in your animal ontology, the fact that penguins swim with the help of their wings, whereas ducks swim with the help of their feet, and ducks fly with the help of their wings. Formulate 2-3 OWL axioms in Manchester Syntax that describe the above behaviour (minor syntax errors will be ignored).

(6 marks)

3. This question is on ontology, OWL and Description Logic.

✧ How do you understand "An ontology being equivalent to a Description Logic knowledge base"?

(3 marks)

✧ Description logics (DLs) are a family of knowledge representation languages that are widely used in ontology development. An important practical reason for this is that they provide one of the main underpinnings for the Web Ontology Language OWL as standardised by the World Wide Web Consortium (W3C). However, DLs have been used in knowledge representation long before the advent of ontologies in the context of the Semantic Web, tracing back to first DL modelling languages in the mid 1980s.

As their name suggests, DLs are logics (in fact they are decidable fragments of first-order logic), and as such they are equipped with a formal semantics: a precise specification of the meaning of DL ontologies. This formal semantics allows humans and computer systems to exchange DL ontologies without ambiguity as to their intended meaning, and also makes it possible to use logical deduction to infer additional information from the facts stated explicitly in an ontology - an important feature that distinguishes DLs from other modelling languages such as UML.

The capability of inferring additional knowledge increases the modelling power of DLs but it also requires some understanding on the side of the modeller and, above all, good tool support for computing the conclusions. The computation of inferences is called reasoning and an important goal of DL language design has been to ensure that reasoning algorithms of good performance are available.

What is the best balance between expressivity of the language and complexity of reasoning?

(3 marks)

4. This question is on syntax, semantics and OWL API.

✧ Consider the interpretation I defined by

$\Delta^I = \{a,b,c\}$;
$A^I = \{a,b\}$, $B^I = \{c\}$, $r^I = \{(a,b), (a,c), (b,b)\}$.

Determine the following sets:

- $(\exists r.A)^I$
- $(\forall r.A)^I$
- $(\exists r.B)^I$
- $(\forall r.B)^I$
- $(\exists r.top)^I$
- $(\forall r.top)^I$

(6 marks)

5. This question is on tableaux algorithm.

✧ Consider the ALC-concept

```
C = (R some (A or (R some false)) and (only R (not A))
```

Apply the ALC-tableaux algorithm to the concept C to determine whether C is satisfiable or not. In your answer, show how the completion rules →and, →or, →exists and →forall are applied step by step to the constraint system x:C. If C is satisfiable, construct an interpretation I satisfying C.

(5 marks)

✧ Consider the following TBox $\mathcal{T}$:

$$\text{Vegan} \equiv \text{Person and } \forall \text{eats.Plant}$$

$$\text{Vegetarian} \equiv \text{Person and } \forall \text{eats.(Plant or Dairy)}$$

We want to know if $\mathcal{T} \models \text{Vegan} \sqsubseteq \text{Vegetarian}$.

(5 marks)

6. This question is on ontology-based query answering.

✧ Consider the TBox T containing

- $\text{Cat} \sqcap \text{Dog} \sqsubseteq \bot$
- $\text{Mammal} \equiv \text{Cat or Dog}$
- $\text{Mammal} \sqsubseteq \exists \text{breatheWith.Lung}$
- $\text{Mammal} \sqsubseteq \forall \text{breatheWith.Lung}$
- $\text{True} \sqsubseteq \forall \text{breatheWith.Lung}$

Consider the ABox A containing

- `Cat(mimi)`
- `¬Dog(gigi)`
- `Mammal(gigi)`
- `breatheWith(dudu, abc)`

Recall that the answers to Boolean queries given by database instances, and knowledge bases are "Yes", "No", "Don't know".

Given the answers given by

- the database instance I_A corresponding to A;
- the knowledge base (T, A),

to the following Boolean queries:

- `Cat(gigi)`
- `Dog(mimi)`
- `Mammal(dudu)`
- `Lung(abc)`

Give a brief explanation for each answer.

(6 marks)

✧ Why is data complexity regarded as a more informative measure of the complexity of query answering in relational databases than combined complexity?

(3 marks)

✧ Consider the TBox $\mathcal{T}$ containing:

$$\text{true} \sqsubseteq \text{red or green}, \quad \text{red and } \exists r.\text{green} \sqsubseteq \text{false}, \quad \text{green and } \exists r.\text{red} \sqsubseteq \text{clash}$$

ABox $\mathcal{A}$ containing:

$$r(0,1),\ r(1,2),\ r(2,3),\ r(3,0)$$

What is the answer of $(\mathcal{T}, \mathcal{A})$ to the Boolean query $\exists x.\text{clash}(x)$? Explain your answer.

(4 marks)

END OF EXAMINATION